

Article

Baseline Methods for Bayesian Inference in Gumbel Distribution

Jacinto Martín ^{1,†} , María Isabel Parra ^{1,†} , Mario Martínez Pizarro ^{2,†}  and Eva L. Sanjuán ^{3,*,†} 

¹ Departamento de Matemáticas, Facultad de Ciencias, Universidad de Extremadura, 06006 Badajoz, Spain; jrmartin@unex.es (J.M.); mipa@unex.es (M.I.P.)

² Departamento de Matemáticas, Facultad de Veterinaria, Universidad de Extremadura, 10003 Cáceres, Spain; mariomp@unex.es

³ Departamento de Matemáticas, Centro Universitario de Mérida, Universidad de Extremadura, 06800 Mérida, Spain

* Correspondence: etlopez@unex.es

† The authors contributed equally to this work.

Received: 28 September 2020; Accepted: 4 November 2020; Published: 7 November 2020



Abstract: Usual estimation methods for the parameters of extreme value distributions only employ a small part of the observation values. When block maxima values are considered, many data are discarded, and therefore a lot of information is wasted. We develop a model to seize the whole data available in an extreme value framework. The key is to take advantage of the existing relation between the baseline parameters and the parameters of the block maxima distribution. We propose two methods to perform Bayesian estimation. Baseline distribution method (BDM) consists in computing estimations for the baseline parameters with all the data, and then making a transformation to compute estimations for the block maxima parameters. Improved baseline method (IBDM) is a refinement of the initial idea, with the aim of assigning more importance to the block maxima data than to the baseline values, performed by applying BDM to develop an improved prior distribution. We compare empirically these new methods with the Standard Bayesian analysis with non-informative prior, considering three baseline distributions that lead to a Gumbel extreme distribution, namely Gumbel, Exponential and Normal, by a broad simulation study.

Keywords: Bayesian inference; highly informative prior; Gumbel distribution; small dataset

1. Introduction

Extreme value theory (EVT) is a widely used statistical tool for modeling and forecasting the distributions which arise when we study events that are more extreme than any previously observed. Examples of these situations are natural rare events in climatology or hydrology, such as floods, earthquakes, climate changes, etc. Therefore, EVT is employed in several scientific fields, to model and predict extreme events of temperature [1–3], precipitation [4–10] and solar climatology [11–13], as well as in the engineering industry to study important malfunctions (e.g., [14], finance (study of financial crisis), insurance (for very large claims due to catastrophic events) and environmental science (concentration of pollution in the air).

Overall fitting method tries to fit all the historical data to several theoretical distributions and then choose the best one, according to certain criteria. However, since the number of extreme observations is usually scarce, overall fitting works well in the central distribution area, but it can poorly represent the tail area.

In EVT, there are two main approaches, block maxima (BM) method and peak-over-threshold (POT), which are differentiated by the way each model classifies the observations considered as extreme events and then uses them in the data analysis process. In the POT method, the extreme data are the ones over a certain threshold, while, for BM method, data are divided into blocks of equal size, and the maximum datum for each block is selected. BM method is preferable to POT when the only available information is block data, when seasonal periodicity is given or the block periods appear naturally, such as in studies for temperature, precipitation and solar climatology. The biggest challenge in BM method is deciding the size of the blocks when they are not obvious.

The Gumbel distribution plays an important role in extreme value analysis as model for block maxima, because it is appropriate to describe extreme events from distributions such as Normal, Exponential and Gumbel distributions [15]. To estimate maximum data distribution, both frequentist and Bayesian approaches have been developed [16,17]. However, the knowledge of physical constraints, the historical evidence of data behavior or previous assessments might be an extremely important matter for the adjustment of the data, particularly when they are not completely representative and further information is required. This fact leads to the use of Bayesian inference to address the extreme value estimation [18].

Practical use of Bayesian estimation is often associated with difficulties to choose prior information and prior distribution for the parameters of the extreme values distribution [19]. To fix this problem, several alternatives have been proposed, either by focusing exclusively on the selection of the prior density [20,21] or by improving the algorithm for the estimation of the parameters [22]. However, the lack of information still seems to be the weakness when referring to extreme value inference.

Examples for its application are the modeling of annual rainfall maximum intensities [23], the estimation of the probability of exceedence of future flood discharge [24] and the forecasting of the extremes of the price distribution [25]. Some of these works are focused on the construction of informative priors of the parameters for which data can provide little information. Despite these previous efforts, it is well understood that some constraints to quantify qualitative knowledge always appear when referring to construct informative priors.

Therefore, this paper focuses on techniques to employ all the available data in order to elicit a highly informative prior distribution. We consider several distributions that lead to a Gumbel extreme distribution. The key is to take advantage of the existing relation between the baseline parameters and the parameters of the block maxima distribution. The use of the entire dataset, instead of the selected block maximum data, results to be adequate and it is advisable when dealing with very shortened available data.

We employ MCMC techniques, concretely a Metropolis–Hastings algorithm. Several statistical analyses are performed to test the validity of our method and check its enhancements in relation to the standard Bayesian analysis without this information.

2. Domains of Attraction of Gumbel Distribution

As is well known, the BM approach consists on dividing the observation period into non overlapping periods of equal size and select the maximum observation in each period. Given a sequence of i.i.d. random variables Y_1, Y_2, \dots, Y_m with common distribution function F , and given a fixed $k \in \mathbb{N}$ (block size), we define the block maxima

$$X_i = \max_{(i-1)k < j \leq ik} Y_j, \quad i = 1, 2, \dots, n. \quad (1)$$

Hence, the total observations, $m = k \times n$, are divided into n blocks of size k . The extreme values depend upon the full sample space from which they have been drawn through its shape and size. Therefore, extremes variate according to the initial distribution and sample size ([26]). Then, the cumulative distribution function of X_i is

$$P(X_i \leq x) = P(Y_{(i-1)k+1} \leq x, \dots, Y_{ik} \leq x) = \prod_{j=(i-1)k+1}^{ik} P(Y_j \leq x) = F(x)^k \tag{2}$$

This result depends on our knowledge of F , in which we could be lacking. Therefore, it is useful to consider the asymptotic distribution.

According to the Gnedenko [27] and Fisher and Tippett [28] theorems, the asymptotic distribution of block maxima of random i.i.d. variables can be approximated by a generalized extreme value distribution, with distribution function

$$GEV(x; \xi, \mu, \sigma) = \exp \left\{ - \left(1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right)^{-1/\xi} \right\} \tag{3}$$

with $\xi, \mu \in \mathbb{R}$, $\sigma > 0$ and $1 + \xi \left(\frac{x - \mu}{\sigma} \right) > 0$.

When $\xi = 0$, the right-hand side of Equation (3) is interpreted as

$$G(x; \mu, \sigma) = \exp \left\{ - \exp \left\{ - \left(\frac{x - \mu}{\sigma} \right) \right\} \right\} \tag{4}$$

and it is called Gumbel distribution with parameters μ (location) and $\sigma > 0$ (scale).

Definition 1. We say that the distribution function F is in the domain of attraction of a extreme value Gumbel distribution when there exist sequences $\{a_k\}$ and $\{b_k\}$, with $a_k > 0, b_k \in \mathbb{R}$ such that

$$\lim_{k \rightarrow \infty} F^k(a_k x + b_k) = G(x), \quad x \in \mathbb{R} \tag{5}$$

Sequences $\{a_k\}$ and $\{b_k\}$ are called normalizing constants. We usually call the distribution F baseline or underlying distribution. Normalizing constants correspond to the parameters of scale and location of the limit Gumbel distribution, therefore they allow us to establish a relation between this distribution and the baseline distribution.

Moreover, Ferreira and de Haan [29] showed theoretical results which allow determining the normalizing constants for many baseline distributions in the domain of attraction of a Gumbel distribution:

Theorem 1. When F belongs to the domain of attraction of a Gumbel distribution, there is a positive function h that verifies

$$a_k = h(b_k), \quad b_k = F^{-1}(1 - k^{-1}), \quad \forall k. \tag{6}$$

To determine function h , the following condition is very useful.

Theorem 2 (Von-Mises condition). When $F''(x)$ and $F'(x)$ exist, and F' is positive for all x belonging to a neighborhood at the left of x^* (right endpoint of F), and

$$\lim_{t \rightarrow x^*} \left(\frac{1 - F}{F'} \right)'(t) = 0, \tag{7}$$

or equivalently

$$\lim_{t \rightarrow x^*} \frac{(1 - F(t)) \cdot F''(t)}{(F'(t))^2} = -1, \tag{8}$$

then F belongs to the domain of attraction of the Gumbel distribution. In this case, function h is determined by

$$h(t) = \frac{1 - F(t)}{F'(t)} \tag{9}$$

Distributions whose tails decrease exponentially produce a Gumbel distribution when taking the block maxima. Besides the Exponential distribution, the class of distributions which belong to the domain of attraction of the Gumbel includes the Normal distribution, and many others, such as Log-normal, Gamma, Rayleigh, Gumbel, etc.

We also use the following result:

Proposition 1. *If X is a random variable belonging to the domain of attraction of a Gumbel distribution, then $Y = \mu + \sigma X$ also belongs to the same domain of attraction. The normalization constants are:*

$$\tilde{a}_k = \sigma a_k, \quad \tilde{b}_k = \mu + \sigma b_k. \tag{10}$$

where a_k and b_k are the normalization constants of X .

2.1. Gumbel Baseline Distribution

If $Y \sim \mathcal{G}(\mu, \sigma)$, then block maxima distribution of size k , denoted by X , is also a Gumbel distribution, because

$$F(x)^k = \exp \left\{ - \exp \left(- \frac{x - \mu}{\sigma} \right) + k \right\} = \exp \left\{ - \exp \left(- \frac{x - (\mu + \sigma \ln k)}{\sigma} \right) \right\}, \tag{11}$$

therefore $X \sim \mathcal{G}(\mu + \sigma \ln k, \sigma)$.

2.2. Exponential Baseline Distribution

Let $Y \sim \text{Exp}(\lambda)$ with distribution function

$$F(y) = 1 - e^{-\lambda y}, \quad y \geq 0, \tag{12}$$

Exponential distribution belongs to the domain of attraction of the Gumbel, with $h(t) = \lambda^{-1}$. As $F^{-1}(u) = \lambda^{-1} \ln(1 - u)$, the normalization constants are

$$a_k = \lambda^{-1}, \quad b_k = \lambda^{-1} \ln k, \tag{13}$$

and they settle a relation that allow us to make estimations for Gumbel limit distribution, when there is an exponential baseline distribution for k big enough.

2.3. Normal Baseline Distribution

When the baseline distribution is a Standard Normal distribution, normalizing constants can be computed, making use of asymptotic limit and results showed before.

Let $\mathcal{Z} \sim \mathcal{N}(0, 1)$, with distribution function F and density function f . It is easy to show that F verifies von Mises condition (8):

$$\begin{aligned} \lim_{t \rightarrow x^*} \frac{(1 - F(t)) \cdot F''(t)}{(F'(t))^2} &= \lim_{t \rightarrow x^*} \frac{-(1 - F(t)) \cdot f(t)t}{(f(t))^2} = \lim_{t \rightarrow x^*} \frac{-(1 - F(t)) \cdot t}{f(t)} \\ &= \lim_{t \rightarrow x^*} \frac{-(1 - F(t)) + f(t) \cdot t}{f'(t)} = \lim_{t \rightarrow x^*} \frac{f(t) \cdot t}{-t \cdot f(t)} = -1 \end{aligned}$$

using L'Hôpital and noticing that $f'(t) = -t \cdot f(t)$. Therefore, $1 - F(t) \approx f(t) \cdot t^{-1}$, and, consequently, the function h defined as (9) verifies

$$\lim_{t \rightarrow x^*} h(t) = t^{-1}$$

Besides, by (6), $F(b_k) = 1 - k^{-1}$. Therefore, $\ln(1 - F(b_k)) = -\ln k$, or

$$\ln f(b_k) - \ln b_k = -\ln k,$$

so

$$b_k^2 + \ln(2\pi) + 2 \ln b_k = 2 \ln k. \tag{14}$$

Defining the function $g(b_k) = b_k^2 + \ln(2\pi) + 2 \ln b_k - 2 \ln k$, and developing its Taylor series around $(2 \ln k)^{1/2}$, we obtain

$$\begin{aligned} g(b_k) &= g\left((2 \ln k)^{1/2}\right) + g'\left((2 \ln k)^{1/2}\right) \cdot \left(b_k - (2 \ln k)^{1/2}\right) + \mathcal{O}\left((2 \ln k)^{1/2}\right) \\ &= [\ln \ln k + \ln 4\pi] + 2 \left[(2 \ln k)^{1/2} + (2 \ln k)^{-1/2} \right] \cdot \left(b_k - (2 \ln k)^{1/2}\right) \\ &\quad + \mathcal{O}\left((2 \ln k)^{1/2}\right), \end{aligned}$$

so, as $g(b_k) = 0$, for k big enough

$$b_k = (2 \ln k)^{1/2} - 2^{-1} (2 \ln k)^{-1/2} [\ln(\ln k) + \ln(4\pi)]. \tag{15}$$

In addition, as $a_k = h(b_k) \approx b_k^{-1}$ and

$$b_k^{-1} = \frac{1}{(2 \ln k)^{1/2} - 2^{-1} (2 \ln k)^{-1/2} [\ln(\ln k) + \ln(4\pi)]} \approx (2 \ln k)^{-1/2} \tag{16}$$

a_k can be taken as

$$a_k \approx (2 \ln k)^{-1/2}. \tag{17}$$

Besides, as a consequence of Theorem 10, if $Y \sim \mathcal{N}(\mu, \sigma)$, for k big enough, the normalization constants are, approximately,

$$a_k = \sigma (2 \ln k)^{-1/2}, \quad b_k = \mu + \sigma \left[(2 \ln k)^{1/2} - 2^{-1} (2 \ln k)^{-1/2} [\ln \ln k + \ln(4\pi)] \right]. \tag{18}$$

2.4. Other Baseline Distributions

This way of working can be extended to other baseline distributions, whose block maxima limit is also a Gumbel, by using existing relations between baseline and limit parameters. In Table 1, normalization constants computed for the most employed distribution functions in the domain of attraction of the Gumbel distribution are shown. Constants a_N and b_N are the normalization constants for Standard Normal distribution, given by (15) and (17), respectively.

Table 1. Normalization constants computed for the most employed distribution function.

Baseline Distribution F	a_k	b_k
Exponential (λ)	λ^{-1}	$\lambda^{-1} \ln k$
Gamma (α, β)	β	$\beta [\ln k + (\alpha - 1) \ln \ln k - \ln \Gamma(\alpha)]$
Gumbel (μ, σ)	σ	$\mu + \sigma \ln k$
Log-Normal (μ, σ)	$a_N \sigma e^{\mu + \sigma b_N}$	$e^{\mu + \sigma b_N}$
Normal (μ, σ)	$\sigma (2 \ln k)^{-1/2}$	$\mu + \sigma \left[(2 \ln k)^{1/2} - \frac{\ln \ln k + \ln 4\pi}{2 (2 \ln k)^{1/2}} \right]$
Rayleigh (σ)	$\sigma (2 \ln k)^{-1/2}$	$\sigma (2 \ln k)^{1/2}$

3. Bayesian Estimation Methods

3.1. Classical Bayesian Estimation for the Gumbel Distribution

To make statistical inferences based on the Bayesian framework, after assuming a prior density for the parameters, $\pi(\theta)$, and combining this distribution with the information brought by the data which are quantified by the likelihood function, $L(\theta|x)$, the posterior density function of the parameters can be determined as

$$\pi(\theta|x) \propto L(\theta|x)\pi(\theta) \quad (19)$$

The remaining of the inference process is fulfilled based on the obtained posterior distribution.

The likelihood function for $\theta = (\mu, \sigma)$, given the random sample $\mathbf{x} = (x_1, \dots, x_n)$ from a Gumbel(μ, σ) distribution, with density function given by

$$f(x|\mu, \sigma) = \frac{1}{\sigma} \exp\left(-\exp\left(-\frac{x-\mu}{\sigma}\right) - \frac{x-\mu}{\sigma}\right) \quad (20)$$

where $\mu \in \mathbb{R}, \sigma \in \mathbb{R}_+^*$, is

$$L(\mu, \sigma|\mathbf{x}) = \frac{1}{\sigma^n} \exp(\Delta), \quad (21)$$

with

$$\Delta = -\sum_{i=1}^n \exp\left(-\frac{x_i-\mu}{\sigma}\right) - \sum_{i=1}^n \frac{x_i-\mu}{\sigma}. \quad (22)$$

In the case of the Gumbel distribution, Rostami and Adam [21] selected eighteen pairs of priors based on the parameters' characteristics, assumed independence, and compared the posterior estimations by applying Metropolis–Hastings (MH) algorithm, concluding that the combination of Gumbel and Rayleigh is the most productive pair of priors for this model. For fixed initial hyper-parameters $\mu_0, \sigma_0, \lambda_0$

$$\begin{aligned} \pi(\mu) &\propto \exp\left(-\exp\left(-\frac{\mu-\mu_0}{\sigma_0}\right) - \frac{\mu-\mu_0}{\sigma_0}\right) \\ \pi(\sigma) &\propto \sigma \exp\left(-\frac{\sigma^2}{2\lambda_0^2}\right). \end{aligned} \quad (23)$$

The posterior distribution is

$$\pi(\mu, \sigma|\mathbf{x}) \propto \frac{1}{\sigma^{n-1}} \exp\left(\Delta - \exp\left(-\frac{\mu-\mu_0}{\sigma_0}\right) - \frac{\mu-\mu_0}{\sigma_0} - \frac{\sigma^2}{2\lambda_0^2}\right), \quad (24)$$

and conditional posterior distributions are

$$\begin{aligned} \pi(\mu|\sigma, \mathbf{x}) &\propto \exp\left(\Delta - \exp\left(-\frac{\mu-\mu_0}{\sigma_0}\right) - \frac{\mu-\mu_0}{\sigma_0}\right) \\ \pi(\sigma|\mu, \mathbf{x}) &\propto \frac{1}{\sigma^{n-1}} \exp\left(\Delta - \frac{\sigma^2}{2\lambda_0^2}\right) \end{aligned} \quad (25)$$

Then, an MCMC method is applied through the MH algorithm.

1. Draw a starting sample $(\mu^{(0)}, \sigma^{(0)})$ from starting distributions, $\pi(\mu), \pi(\sigma)$, respectively, given by Equation (23).
2. For $j = 0, 1, \dots$, given the chain is currently at $\mu^{(j)}, \sigma^{(j)}$,

- Sample candidates μ^*, σ^* for the next sample from a proposal distribution,

$$\mu^* \sim \mathcal{N}(\mu^{(j)}, v_\mu) \text{ and } \sigma^* \sim \mathcal{N}(\sigma^{(j)}, v_\sigma)$$

- Calculate the ratios

$$r_\mu = \frac{\pi(\mu^* | \sigma^{(j)}, \mathbf{x})}{\pi(\mu^{(j)} | \sigma^{(j)}, \mathbf{x})}, \quad r_\sigma = \frac{\pi(\sigma^* | \mu^{(j)}, \mathbf{x})}{\pi(\sigma^{(j)} | \mu^{(j)}, \mathbf{x})} \quad (26)$$

- Set

$$\mu^{(j+1)} = \begin{cases} \mu^*, & \text{with probability } \min\{1, r_\mu\} \\ \mu^{(j)}, & \text{otherwise} \end{cases} \quad (27)$$

$$\sigma^{(j+1)} = \begin{cases} \sigma^*, & \text{with probability } \min\{1, r_\sigma\} \\ \sigma^{(j)}, & \text{otherwise} \end{cases} \quad (28)$$

3. Iterate the former procedure. Notice that

$$\begin{aligned} r_\mu &= \exp \left\{ \frac{n}{\sigma^{(j)}} (\mu^* - \mu^{(j)}) + \frac{\mu^{(j)} - \mu^*}{\sigma_0} + \exp \left(-\frac{\mu^{(j)} - \mu_0}{\sigma_0} \right) - \exp \left(-\frac{\mu^* - \mu_0}{\sigma_0} \right) \right. \\ &\quad \left. + \sum_{i=1}^n \left[\exp \left(-\frac{x_i - \mu^{(j)}}{\sigma^{(j)}} \right) - \exp \left(-\frac{x_i - \mu^*}{\sigma^{(j)}} \right) \right] \right\} \\ r_\sigma &= \left(\frac{\sigma^{(j)}}{\sigma^*} \right)^{n-1} \exp \left\{ \frac{(\sigma^{(j)})^2 - (\sigma^*)^2}{2\lambda_0^2} + \left(\frac{1}{\sigma^{(j)}} - \frac{1}{\sigma^*} \right) \sum_{i=1}^n (x_i - \mu^{(j)}) \right. \\ &\quad \left. + \sum_{i=1}^n \left[\exp \left(-\frac{x_i - \mu^{(j)}}{\sigma^{(j)}} \right) - \exp \left(-\frac{x_i - \mu^{(j)}}{\sigma^*} \right) \right] \right\}. \end{aligned}$$

Therefore, we obtain a Markov chain that converges to the posterior distributions for the parameters μ and σ . We call this method Classical Metropolis–Hastings method (MHM).

3.2. Baseline Distribution Method

In Baseline distribution method (BDM), we take all the information available to determine posterior baseline distribution. We denote $\mathcal{B}(\theta)$ as the baseline distribution with parameter vector θ .

Then, we can apply Bayesian inference procedures to estimate the posterior distribution of the baseline distribution, denoted by $\pi_b(\theta | \mathbf{y})$ and, therefore, to obtain estimations for the parameters of the baseline distribution θ , with all the data provided by \mathbf{y} .

Afterwards, making the transformation given by the relations we obtained in previous section, we can obtain new estimations for the parameters of block maxima distribution, which is the Gumbel in this case. We explain the procedure for the three baseline distributions considered in this paper: Gumbel, Exponential and Normal distribution.

3.2.1. Gumbel Baseline Distribution

When the baseline distribution $Y \sim \mathcal{G}(\mu_b, \sigma_b)$, it is known that the limit distribution $X \sim \mathcal{G}(\mu_b + \sigma_b \ln k, \sigma_b)$. Therefore, MH algorithm can be applied to the whole dataset, \mathbf{y} , to find estimations for μ_b and σ_b . Afterwards, we make the adequate transformation to compute estimations for the parameters of X .

3.2.2. Exponential Baseline Distribution

When the baseline distribution $Y \sim \text{Exp}(\lambda_b)$, we consider a Gamma distribution with parameters α_0 and β_0 as prior distribution

$$\pi(\lambda_b) \propto \lambda_b^{\alpha_0 - 1} \exp(-\beta_0 \lambda_b). \quad (29)$$

Therefore, the posterior distribution is

$$\pi(\lambda_b | \mathbf{y}) \sim \Gamma\left(\alpha_0 + m, \beta_0 + \sum_{j=1}^m y_j\right), \tag{30}$$

thus Gibbs algorithm can be employed to generate samples of posterior distribution $\pi(\lambda_b | \mathbf{y})$. Once the estimation of λ_b is obtained, k th power of the distribution function will be the estimation for block maxima distribution function of size k .

3.2.3. Normal Baseline Distribution

Finally, when the baseline distribution $Y \sim \mathcal{N}(\mu_b, \sigma_b)$, we employ Normal and inverse Gamma prior distributions

$$\begin{aligned} \pi(\mu_b) &\propto \exp\left(-\frac{\sigma_0}{2\sigma_b^2}(\mu - \mu_0)^2\right), \\ \pi(\sigma_b^2) &\propto \left(\frac{1}{\sigma_b^2}\right)^{\alpha_0-1} \exp\left(-\frac{\beta_0}{\sigma_b^2}\right). \end{aligned} \tag{31}$$

Therefore, posterior distributions are

$$\begin{aligned} \pi(\mu_b | \mathbf{y}, \sigma_b^2) &\sim \mathcal{N}\left(\frac{\sigma_0\mu_0 + \sum_{j=1}^m y_j}{\sigma_0 + m}, \frac{\sigma_b^2}{\sigma_0 + m}\right), \\ \pi(\sigma_b^2 | \mathbf{y}, \mu_b) &\sim \text{Inv}\Gamma\left(\frac{m}{2} + \alpha_0, \beta_0 + \frac{1}{2}\sum_{j=1}^m (y_j - \mu_b)^2\right), \end{aligned} \tag{32}$$

and we can employ Gibbs algorithm to generate samples of posterior distribution, and, afterwards, the k th power of the distribution function, as in the previous case.

3.3. Improved Baseline Distribution Method

Finally, we propose a new method, called Improved Baseline distribution method (IBDM), to import the highly informative baseline parameters into the Bayesian inference procedure. Here, we take into account the spirit of classical EVT, which grants more weight to block maxima data than to baseline data.

The method consists on applying BDM to obtain the posterior distribution for the parameters of the baseline distribution $\pi(\boldsymbol{\theta} | \mathbf{y})$, and then uses it to build a prior distribution for the parameters of the Gumbel. Therefore, we have a highly informative prior distribution.

As the priors are highly informative, $\pi(\boldsymbol{\theta}^*) = \pi(\boldsymbol{\theta}^{(j)})$, the ratio in the j th step of MH algorithm is

$$\begin{aligned} r_{\boldsymbol{\theta}} = \frac{L(\mu^*, \sigma^* | \mathbf{x})}{L(\mu^{(j)}, \sigma^{(j)} | \mathbf{x})} &= \left(\frac{\sigma^{(j)}}{\sigma^*}\right)^n \exp\left\{\sum_{i=1}^n \left[\exp\left(-\frac{x_i - \mu^{(j)}}{\sigma^{(j)}}\right) \right. \right. \\ &\quad \left. \left. - \exp\left(-\frac{x_i - \mu^*}{\sigma^*}\right) + \frac{x_i - \mu^{(j)}}{\sigma^{(j)}} - \frac{x_i - \mu^*}{\sigma^*}\right]\right\}. \end{aligned}$$

For every iteration of the algorithm, we first make an iteration of Baseline Distribution method, resulting $\boldsymbol{\theta}_b$ as estimation of the posterior distribution $\pi(\boldsymbol{\theta}_b | \mathbf{y})$. Afterwards, a candidate $\boldsymbol{\theta}^*$ is generated using a Normal distribution $\mathcal{N}(f(\boldsymbol{\theta}_b), \nu_{\boldsymbol{\theta}})$ with the adequate transformation $f(\boldsymbol{\theta}_b)$, given by Equations (11), (13) and (18) in the case of Gumbel, Exponential or Normal baseline distributions, respectively.

Obviously, this method is quite similar to BDM when block size is big and, consequently, there are few maxima data. It approaches the classical Bayesian method as the block size gets smaller (more maxima data).

4. Simulation Study

We made a simulation study for the three distributions analyzed above, which belong to the domain of attraction of the Gumbel distribution: Gumbel, Exponential and Normal.

For each distribution selected (once its parameters are fixed), we generated $m_{ij} = n_i \times k_j$ values, where

- n_i is the number of block maxima, $n_i = 2^i, i = 1, 2, \dots, 7$; and
- k_j is the block size, $k_j = 10^j, j = 1, 2, 3$.

Therefore, the sample sizes vary from 20 to 128,000. Besides, each sequence is replicated 100 times. Consequently, 2100 sequences of random values were generated for each combination of parameters of each baseline distribution.

To guarantee the convergence of the MCMC algorithm, we must be sure that the posterior distribution has been reached. Some proceedings are advisable to be performed.

- Burn-in period: Eliminate the first generated values.
- Take different initial values and select them for each sample.
- Make a thinning to assure lack of autocorrelation.

These proceedings were made using library coda [30] for R software, taking 3000 values for the burn-in period, 50 values for the thinning and selecting initial values for each sample. Finally, to get the posterior distribution for each parameter, a Markov chain of length 10,000 was obtained. Therefore, 53,000 iterations were made for each sequence.

There are some common features for the baseline distributions considered when comparing the three methods MHM, BDM and IBDM.

1. To choose an estimator for the parameters, we compared mean- and median-based estimations. They were reasonably similar, due to the high symmetry of posterior distributions. Therefore, we chose the mean of the posterior distribution to make estimations of the parameters.
2. MHM usually provides high skewed estimations for the posterior distributions. BDM is the method that shows less skewness.
3. BDM is the method that offers estimations for posterior distribution with less variability. IBDM provides higher variability, but we must keep in mind that this method stresses the importance of extreme values, therefore more variability is expectable than the one provided by BDM. The method with highest variability is MHM.
4. The election of the most suitable method also depends on the characteristics of the problem. When block maxima data are very similar to the baseline distribution, BDM provides the best estimations and the lowest measures of error. On the contrary, when extreme data differ from baseline data, IBDM offers the lowest errors. IBDM is the most stable method: regardless of the differences between extreme data and baseline data, it provides reasonably good measures of error.

4.1. Gumbel Baseline Distribution

We considered the baseline $\mathcal{G}(\mu_b, \sigma_b)$ distribution. As the localization parameter has no effect on the variability of data, its value was fixed as $\mu_b = 0$ for easiness. Scale parameter does affect variability, so we considered values $\sigma_b = 2^j, j = -2, -1, 0, 1, 2$.

One important point of the simulation study is to observe how the estimation of the parameters vary for a fixed block size as the number of block maxima n (the amount of information we have) is

changing. Regardless of the chosen method, as n increases, variability and skewness decreases. However, for small values of n , BDM and IBDM provide more concentrated and less skewed distributions than the ones offered by MHM. We can appreciate this in Figure 1, where the probability density functions (pdf) are shown for the 100 estimations of the mean for the parameters μ (left) and σ (right) for block maxima Gumbel distributions, with the three methods. The baseline distribution is $\mathcal{G}(0, 4)$ and fixed block size $k = 1000$. Therefore, block maxima distribution is $\mathcal{G}(27.63, 4)$ (from (11)). Scales are very different in this charts, due to a better visualization of distributions. For example, for MHM, highest value of the pdf for $\hat{\sigma}$ is around 1.5 (for $n = 128$), but it is over 40 for BDM.

This behavior is shown qualitatively for all the values of the parameters employed in the simulation.

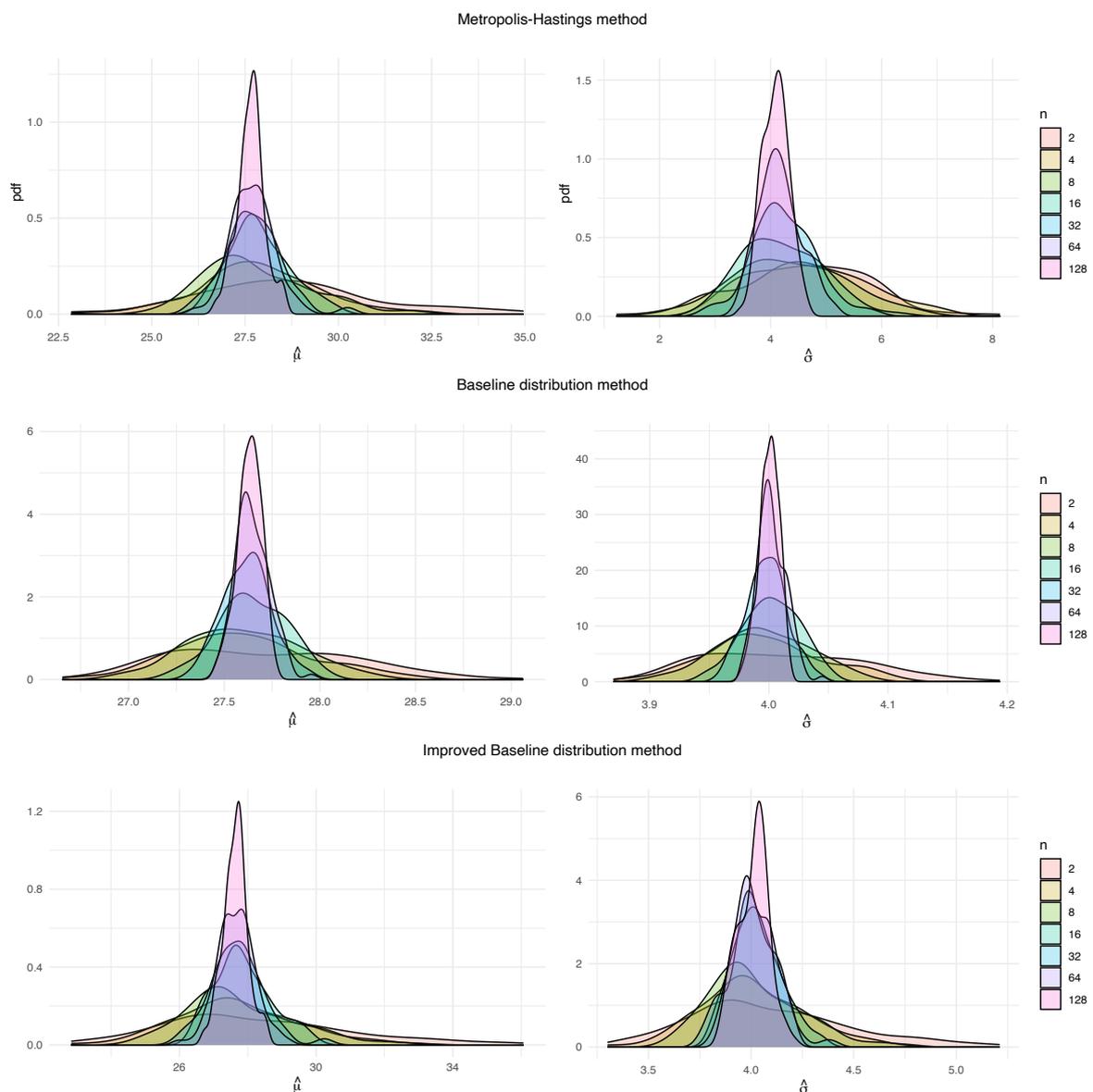


Figure 1. Probability density functions for 100 estimations of the block maxima parameters μ (left) and σ (right), obtained for the three methods, with $k = 1000$ and different values of n , from $\mathcal{G}(0, 4)$ baseline distribution.

To compute measures of error, in the case of Gumbel baseline distribution, we can employ absolute errors $AE_i = |\hat{\theta}_i - \theta|$, where $\hat{\theta}_i$ is the estimation obtained from i th sample and θ is the real value of the estimated parameter. We can then define

- Mean error:

$$ME = \frac{1}{M} \sum_{i=1}^M (\hat{\theta}_i - \theta).$$

- Root mean square error:

$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M (\hat{\theta}_i - \theta)^2}.$$

- Mean absolute error:

$$MAE = \frac{1}{M} \sum_{i=1}^M |\hat{\theta}_i - \theta|,$$

where M is the number of samples.

The three methods provide estimations with low absolute errors AE when the number of maxima n is high, and especially when block size k is high. When both numbers are small, BDM and IBDM get closer estimations and differ from MHM.

In Tables 2 and 3, we show values for ME and RMSE for the estimations of parameters μ and σ , respectively, for some values of k , n and σ_b , for a $\mathcal{G}(0, \sigma_b)$ baseline distribution. We can see that BDM is the method that offers lower values for both measures of error, followed by IBDM. The method that provides highest errors is MHM.

Table 2. ME for μ , with RMSE in brackets, for a baseline distribution $\mathcal{G}(0, \sigma_b)$.

		<i>k</i>									
		10			100			1000			
<i>n</i>		MHM	BDM	IBDM	MHM	BDM	IBDM	MHM	BDM	IBDM	σ_b
2		−0.0879	0.0752	−0.0038	−0.1056	0.0044	−0.0450	−0.0495	0.0061	0.0159	1/4
		(0.2290)	(0.1594)	(0.1848)	(0.2136)	(0.0635)	(0.1658)	(0.1659)	(0.0343)	(0.1782)	
		0.3128	0.3376	0.3350	0.3326	0.1017	0.2866	0.1578	0.0354	0.1137	1
		(0.9906)	(0.6637)	(0.7905)	(0.8535)	(0.3046)	(0.6906)	(0.6573)	(0.1464)	(0.6850)	
	1.6813	1.0156	1.3856	0.2687	−0.0001	0.3466	1.0252	0.0747	0.6297	4	
	(4.6198)	(2.2176)	(2.6874)	(3.1694)	(1.1367)	(2.0640)	(2.6346)	(0.4899)	(2.8024)		
16		0.0040	0.0036	−0.0026	−0.0018	−0.0050	−0.0083	0.0118	−0.0002	0.0014	1/4
		(0.0677)	(0.0477)	(0.0680)	(0.0645)	(0.0260)	(0.0636)	(0.0678)	(0.0119)	(0.0636)	
		0.0712	0.0505	0.0528	0.0454	−0.0033	0.0218	0.0499	0.0006	0.0219	1
		(0.2556)	(0.1966)	(0.2470)	(0.2821)	(0.1019)	(0.2657)	(0.2536)	(0.0441)	(0.2446)	
	0.2096	0.1711	0.2450	0.0625	−0.0010	0.0861	0.2545	0.0178	0.2324	4	
	(1.3466)	(0.8447)	(1.1989)	(1.0148)	(0.3901)	(1.0165)	(0.8414)	(0.1779)	(0.8607)		
128		0.0018	0.0023	0.0012	0.0012	0.0001	0.0006	−0.0038	−0.0005	−0.0045	1/4
		(0.0225)	(0.0170)	(0.0224)	(0.0233)	(0.0088)	(0.0230)	(0.0208)	(0.0039)	(0.0206)	
		0.0258	0.0155	0.0233	0.0070	0.0006	0.0041	−0.0037	0.0004	−0.0067	1
		(0.0992)	(0.0650)	(0.0982)	(0.0958)	(0.0355)	(0.0953)	(0.0944)	(0.0170)	(0.0948)	
	−0.0077	0.0098	0.0002	0.0070	0.0007	0.0052	0.0215	0.0001	0.0097	4	
	(0.3712)	(0.2236)	(0.3623)	(0.3897)	(0.1425)	(0.3879)	(0.3650)	(0.0636)	(0.3632)		

Table 3. ME for σ , with RMSE in brackets, for a baseline distribution $\mathcal{G}(0, \sigma_b)$.

n	k									σ_b
	10			100			1000			
	MHM	BDM	IBDM	MHM	BDM	IBDM	MHM	BDM	IBDM	
2	2.4739	0.0309	0.0258	2.6628	0.0013	0.0048	2.7177	0.0007	0.0024	1/4
	(2.6201)	(0.0584)	(0.0560)	(2.6967)	(0.0119)	(0.0216)	(2.7371)	(0.0045)	(0.0155)	
	2.3399	0.1121	0.1324	2.3384	0.0206	0.0445	2.4139	0.0043	0.0130	1
	(2.4169)	(0.2259)	(0.2591)	(2.3760)	(0.0596)	(0.1085)	(2.4441)	(0.0204)	(0.0774)	
16	0.8238	0.3104	0.4727	0.6607	0.0056	0.0790	0.7284	0.0117	0.0905	4
	(1.6431)	(0.7600)	(0.9749)	(1.1945)	(0.2174)	(0.4192)	(1.2739)	(0.0680)	(0.3978)	
	0.0292	0.0006	0.0031	0.0229	-0.0008	-0.0037	0.0462	0.0000	0.0070	1/4
	(0.0577)	(0.0165)	(0.0279)	(0.0560)	(0.0051)	(0.0264)	(0.0788)	(0.0016)	(0.0293)	
128	0.1520	0.0152	0.0367	0.1339	-0.0003	0.0032	0.1495	0.0001	0.0058	1
	(0.2835)	(0.0604)	(0.1234)	(0.2740)	(0.0198)	(0.0959)	(0.2712)	(0.0059)	(0.0605)	
	0.3237	0.0742	0.1261	0.3159	-0.0010	0.0176	0.2071	0.0023	0.0294	4
	(0.9959)	(0.2826)	(0.4765)	(0.8719)	(0.0755)	(0.2238)	(0.7576)	(0.0237)	(0.1211)	
128	0.0074	0.0009	0.0055	0.0036	0.0001	0.0018	0.0027	-0.0001	0.0009	1/4
	(0.0181)	(0.0058)	(0.0162)	(0.0203)	(0.0017)	(0.0184)	(0.0179)	(0.0005)	(0.0161)	
	0.0300	0.0041	0.0220	0.0313	0.0003	0.0203	0.0138	0.0000	0.0037	1
	(0.0778)	(0.0221)	(0.0693)	(0.0756)	(0.0069)	(0.0616)	(0.0695)	(0.0023)	(0.0518)	
128	0.0557	0.0062	0.0287	0.0415	-0.0020	0.0024	0.0884	-0.0001	0.0123	4
	(0.2423)	(0.0759)	(0.1843)	(0.2805)	(0.0274)	(0.1286)	(0.2522)	(0.0086)	(0.0753)	

4.2. Exponential Baseline Distribution

Assume now we have another baseline distribution F , which is not a Gumbel. Notice that, for methods MHM and IBDM, we are approaching block maxima distribution by a Gumbel. However, for BDM, we employ the k th power of the baseline distribution function. When the baseline distribution is a Gumbel, we know that the k th power is also a Gumbel. However, for another baseline distributions, this is not true.

For this reason, we have to define different measures of error to evaluate the quality of estimations. We compared estimated distribution functions (H) with real ones (F^k) through their mean distance (D), mean absolute distance (AD) and root square distance (RSD). As analytical computation is not possible, we made a Monte-Carlo computation employing sample size $s = 10^4$. Then,

$$D_j = \frac{1}{s} \sum_{i=1}^s (H(x_i; \hat{\theta}_j) - F(x_i; \theta)^k) \tag{33}$$

$$AD_j = \frac{1}{s} \sum_{i=1}^s |H(x_i; \hat{\theta}_j) - F(x_i; \theta)^k| \tag{34}$$

and

$$RSD_j = \sqrt{\frac{1}{s} \sum_{i=1}^s (H(x_i; \hat{\theta}_j) - F(x_i; \theta)^k)^2}, \tag{35}$$

with $j = 1, \dots, M$, where M is the number of samples. $H(x; \hat{\theta})$ denotes the estimated distribution function for block maxima, for the baseline parameter $\hat{\theta}$, which is $\hat{\theta} = \hat{\lambda}_b$ if we have an exponential baseline distribution and $\hat{\theta} = (\hat{\mu}_b, \hat{\sigma}_b)$ for the Normal baseline distribution.

The measures of error were:

- Mean error:

$$ME = \frac{1}{M} \sum_{j=1}^M D_j. \tag{36}$$

- Root mean square error:

$$RMSE = \frac{1}{M} \sum_{j=1}^M RSD_j. \tag{37}$$

- Mean absolute error:

$$MAE = \frac{1}{M} \sum_{j=1}^M AD_j. \tag{38}$$

We considered the baseline $Exp(\lambda_b)$ distribution. In this case, for k big enough, $X \approx \mathcal{G}(\lambda_b^{-1} \ln k, \lambda_b^{-1})$. We took $\lambda_b = 2^j$, with $j = -2, -1, 0, 1, 2$.

As in the Gumbel baseline distribution, MHM shows high skewness when the number of blocks n is very small, compared to IBDM (see Figure 2). In addition, if we compute measures of error, we can see that, for small block sizes, the three methods offer similar values (see Table 4). For bigger values of k , BDM and IBDM provide better results, and usually BDM is the best method.

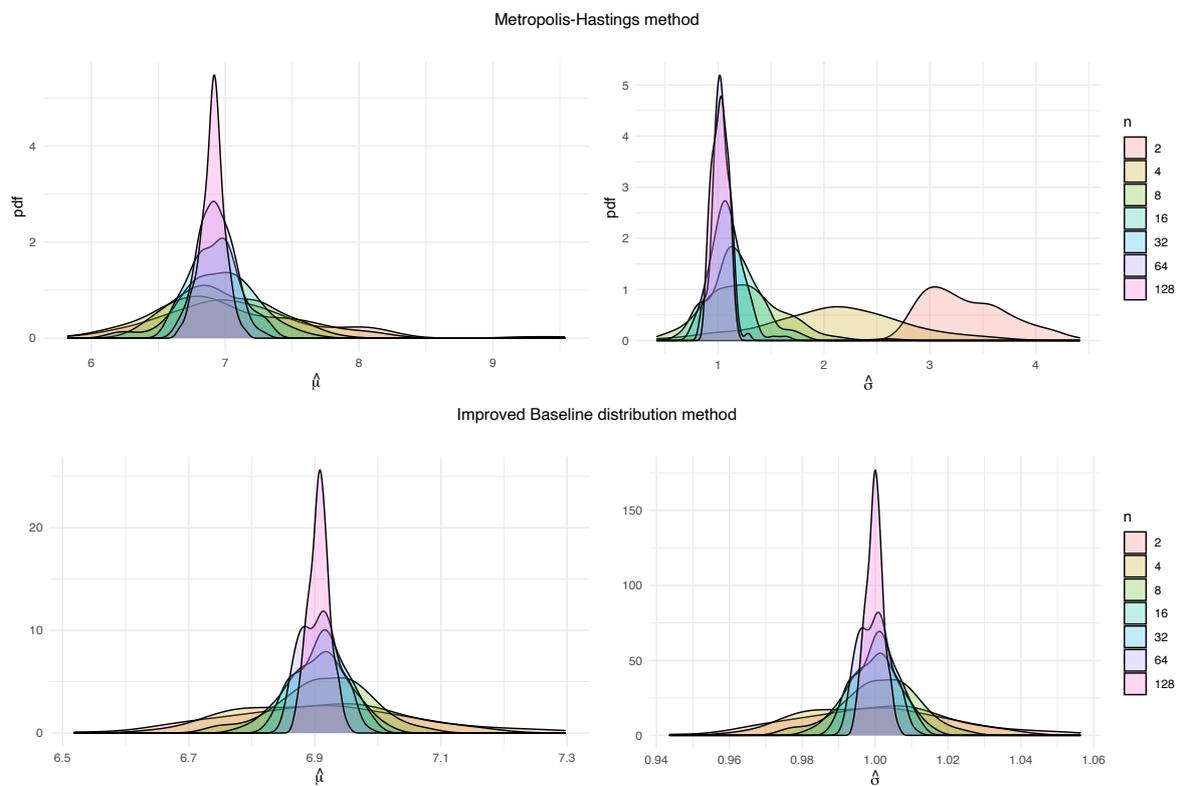


Figure 2. Probability density functions for M estimations of the block maxima parameters μ (left) and σ (right), obtained for the methods MHM and IBDM, with $k = 1000$ and different values of n , from $Exp(1)$ baseline distribution.

Table 4. ME for σ , with RMSE in brackets, for a baseline distribution $Exp(\lambda_b)$.

<i>n</i>	<i>k</i>									λ_b
	10			100			1000			
	MHM	BDM	IBDM	MHM	BDM	IBDM	MHM	BDM	IBDM	
2	−0.0645	0.0287	0.0207	−0.0563	0.0098	0.0061	−0.0652	−0.0004	−0.0011	1/2
	(0.1569)	(0.1297)	(0.1278)	(0.1485)	(0.0880)	(0.0878)	(0.1471)	(0.0339)	(0.0339)	
	−0.0818	0.0157	0.0084	−0.0915	−0.0078	−0.0115	−0.0945	0.0002	−0.0006	1
	(0.2022)	(0.1409)	(0.1382)	(0.2032)	(0.0780)	(0.0782)	(0.2012)	(0.0330)	(0.0330)	
16	−0.0909	0.0299	0.0221	−0.1004	−0.0033	−0.0070	−0.0954	0.0066	0.0058	2
	(0.2442)	(0.1400)	(0.1377)	(0.2448)	(0.0673)	(0.0674)	(0.2437)	(0.0377)	(0.0376)	
	−0.0232	−0.0075	0.0029	−0.0263	−0.0041	−0.0035	−0.0224	−0.0030	−0.0029	1/2
	(0.0725)	(0.0476)	(0.0490)	(0.0743)	(0.0255)	(0.0254)	(0.0705)	(0.0128)	(0.0128)	
128	−0.0156	−0.0006	0.0096	−0.0221	−0.0008	−0.0002	−0.0159	0.0008	0.0008	1
	(0.0680)	(0.0440)	(0.0452)	(0.0832)	(0.0302)	(0.0302)	(0.0750)	(0.0110)	(0.0110)	
	−0.0109	0.0019	0.0120	−0.0230	−0.0027	−0.0021	−0.0220	0.0001	0.0122	2
	(0.0741)	(0.0429)	(0.0455)	(0.0750)	(0.0294)	(0.0293)	(0.0743)	(0.0132)	(0.0132)	
128	−0.0053	−0.0050	0.0076	−0.0003	−0.0007	0.0005	−0.0093	−0.0012	−0.0010	1/2
	(0.0282)	(0.0197)	(0.0229)	(0.0249)	(0.0097)	(0.0097)	(0.0251)	(0.0044)	(0.0044)	
	−0.0039	−0.0057	0.0069	−0.0020	0.0002	0.0013	−0.0007	0.0002	0.0003	1
	(0.0244)	(0.0183)	(0.0218)	(0.0243)	(0.0100)	(0.0101)	(0.0227)	(0.0036)	(0.0036)	
128	−0.0010	−0.0022	0.0103	−0.0002	0.0012	0.0024	−0.0004	0.0012	0.0013	2
	(0.0243)	(0.0153)	(0.0216)	(0.0267)	(0.0113)	(0.0115)	(0.0269)	(0.0044)	(0.0044)	

4.3. Normal Baseline Distribution

Finally, assume the baseline distribution is $\mathcal{N}(\mu_b, \sigma_b)$. As the mean has no effect on the variability of data, its value was fixed as $\mu_b = 0$ for easiness. Standard deviation was taken as $\sigma_b = 2^j$, $j = -2, -1, 0, 1, 2$.

The conclusions we obtain are quite similar to the previous case. We illustrate skewness and variability for MHM employing similar graphs (see Figure 3). Variability is especially pronounced when n is small, and we are estimating σ . In addition, errors are shown in Table 5.

Table 5. ME for σ , with RMSE in brackets, for a baseline distribution $\mathcal{N}(0, \sigma_b)$.

<i>n</i>	<i>k</i>									σ_b
	10			100			1000			
	MHM	BDM	IBDM	MHM	BDM	IBDM	MHM	BDM	IBDM	
2	−0.1097	−0.2932	−0.1942	−0.1120	−0.1183	−0.1204	−0.1101	−0.0132	−0.0319	1/4
	(0.2923)	(0.3274)	(0.2210)	(0.2975)	(0.1356)	(0.1481)	(0.2991)	(0.0443)	(0.0559)	
	−0.1023	−0.0049	−0.0222	−0.1009	0.0078	−0.0077	−0.1053	0.0047	−0.0150	1
	(0.2481)	(0.1139)	(0.1404)	(0.2609)	(0.0718)	(0.0876)	(0.2704)	(0.0394)	(0.0482)	
16	−0.0822	0.0252	−0.0079	−0.0803	0.0079	−0.0103	−0.0824	−0.0045	−0.0227	4
	(0.1643)	(0.1197)	(0.1374)	(0.1790)	(0.0892)	(0.1073)	(0.1876)	(0.0393)	(0.0497)	
	−0.0089	−0.0650	−0.0435	−0.0131	−0.0166	−0.0278	−0.0046	−0.0006	−0.0198	1/4
	(0.0736)	(0.0765)	(0.0644)	(0.0785)	(0.0326)	(0.0500)	(0.0638)	(0.0134)	(0.0283)	
128	−0.0085	0.0024	−0.0030	−0.0092	−0.0042	−0.0164	−0.0109	0.0012	−0.0184	1
	(0.0750)	(0.0546)	(0.0645)	(0.0658)	(0.0282)	(0.0423)	(0.0724)	(0.0144)	(0.0292)	
	−0.0060	0.0033	−0.0049	−0.0171	−0.0022	−0.0185	−0.0028	−0.0006	−0.0194	4
	(0.0679)	(0.0427)	(0.0574)	(0.0719)	(0.0338)	(0.0472)	(0.0675)	(0.0142)	(0.0281)	
128	0.0054	−0.0101	−0.0116	0.0041	−0.0011	−0.0154	0.0021	−0.0008	−0.0207	1/4
	(0.0323)	(0.0188)	(0.0299)	(0.0282)	(0.0116)	(0.0272)	(0.0235)	(0.0044)	(0.0260)	
	0.0090	−0.0002	−0.0038	0.0058	0.0014	−0.0126	0.0016	−0.0002	−0.0204	1
	(0.0336)	(0.0172)	(0.0296)	(0.0301)	(0.0116)	(0.0263)	(0.0249)	(0.0050)	(0.0258)	
128	0.0052	−0.0020	−0.0063	0.0052	−0.0006	−0.0147	0.0020	0.0001	−0.0200	4
	(0.0357)	(0.0188)	(0.0320)	(0.0288)	(0.0116)	(0.0272)	(0.0277)	(0.0048)	(0.0255)	

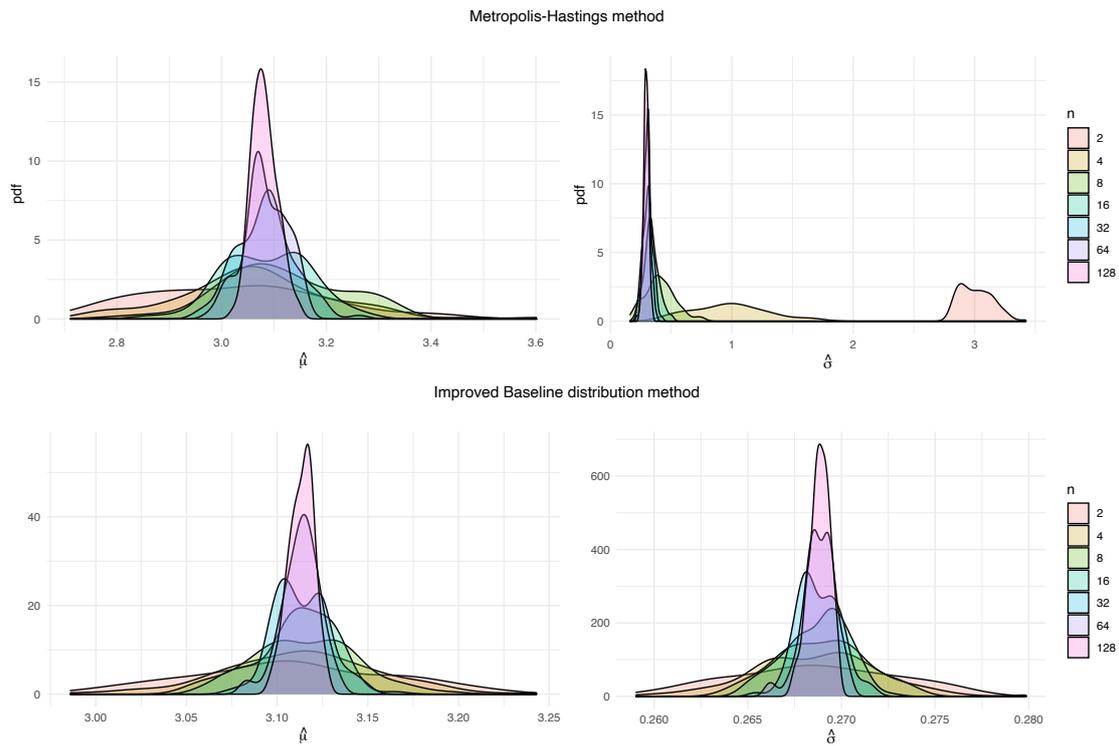


Figure 3. Probability density functions for M estimations of the block maxima parameters μ (left) and σ (right), obtained for the methods MHM and IBDM, with $k = 1000$ and different values of n , from $\mathcal{N}(0, 1)$.

In practical situations, data might not adjust to a concrete distribution and some perturbations (noise) could appear. To get a quick overview of how differences between baseline distribution data and block maxima data can affect the choice of the best method, we simulated a simple situation, when data come from a mixture of normal variables. Concretely,

$$Y = 0.9 \cdot Z + 0.1 \cdot Y_1, \quad Z \sim \mathcal{N}(0, 1), \quad Y_1 \sim \mathcal{N}(1, 1.5).$$

In Figure 4, we can see how MAE vary for the three methods when we vary the number of block maxima n , for a block size $k = 100$. In this case, IBDM offers the lowest errors, because it stresses the importance of extreme data. When the extreme data are scarce, both new methods, BDM and IBDM, improve MHM meaningfully.

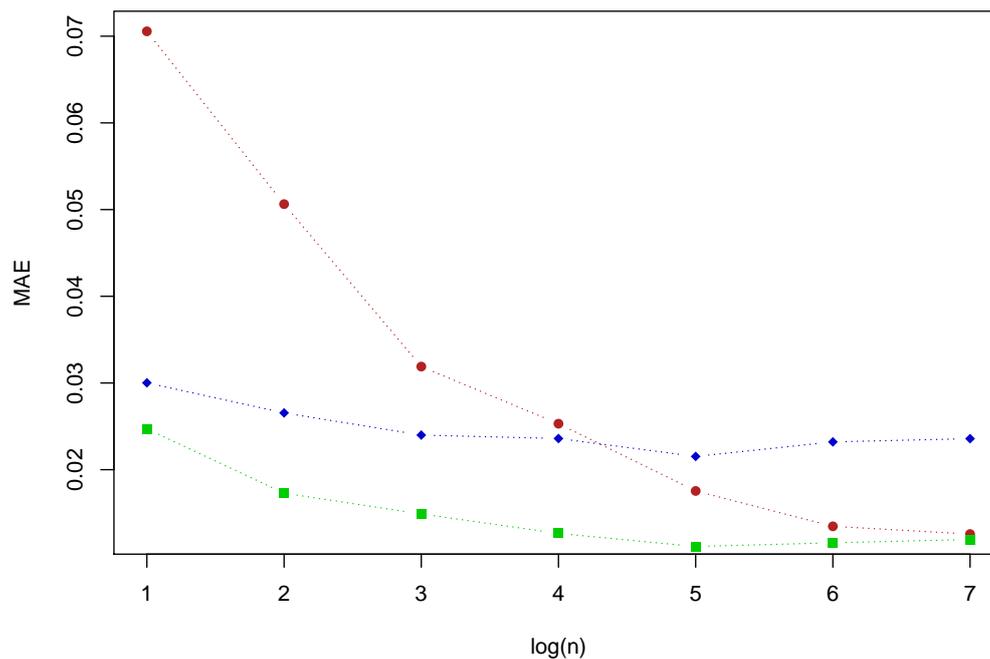


Figure 4. MAE for the three methods MHM (red), BDM (blue) and IBDM (green) with $k = 100$ and different values of n , for the baseline distribution Y .

5. Conclusions

1. One of the most common problems in EVT is estimating the parameters of the distribution, because the data are usually scarce. In this work, we considered the case when block maxima distribution is a Gumbel, and we developed two bayesian methods, BDM and IBDM, to estimate posterior distribution, making use of all the available data of the baseline distribution, not only the block maxima values.
2. The methods were proposed for three baseline distributions, namely Gumbel, Exponential and Normal, but the new strategy can easily be applied to some other baseline distributions, following the relations shown in Table 1.
3. We performed a broad simulation study to compare BDM and IBDM methods to classical Metropolis–Hastings method (MHM). The results are based on numerical studies, but theoretical support still needs to be provided.
4. We obtained that posterior distributions of BDM and IBDM are more concentrated and less skewed than MHM.
5. In general, the results obtained show that the methods which offer lower measures of error are BDM and IBDM, as they leverage all the data. The classical method, MHM, shows the worst results, especially when extreme data are scarce.
6. IBDM is the most stable method: regardless of the differences between extreme data and baseline data, it provides reasonably good measures of error. When the extreme data are scarce, both new methods, BDM and IBDM, improve MHM meaningfully.

Author Contributions: All authors contributed equally to this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by Junta de Extremadura, Consejería de Economía e Infraestructuras FEDER Funds IB16063, GR18108 project from Junta de Extremadura and Project MTM2017-86875-C3-2-R from Ministerio de Economía, Industria y Competitividad de España.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Nogaj, M.; Yiou, P.; Parey, S.; Malek, F.; Naveau, P. Amplitude and frequency of temperature extremes over the North Atlantic region. *Geophys. Res. Lett.* **2006**, *33*. [[CrossRef](#)]
2. Coelho, C.A.S.; Ferro, C.A.T.; Stephenson, D.B.; Steinskog, D.J. Methods for Exploring Spatial and Temporal Variability of Extreme Events in Climate Data. *J. Clim.* **2008**, *21*, 2072–2092. [[CrossRef](#)]
3. Acero, F.J.; Fernández-Fernández, M.I.; Carrasco, V.M.S.; Parey, S.; Hoang, T.T.H.; Dacunha-Castelle, D.; García, J.A. Changes in heat wave characteristics over Extremadura (SW Spain). *Theor. Appl. Climatol.* **2017**, *1–13*. [[CrossRef](#)]
4. García, J.; Gallego, M.C.; Serrano, A.; Vaquero, J. Trends in Block-Seasonal Extreme Rainfall over the Iberian Peninsula in the Second Half of the Twentieth Century. *J. Clim.* **2007**, *20*, 113–130. [[CrossRef](#)]
5. Re, M.; Barros, V.R. Extreme rainfalls in SE South America. *Clim. Chang.* **2009**, *96*, 119–136. [[CrossRef](#)]
6. Acero, F.J.; García, J.A.; Gallego, M.C. Peaks-over-Threshold Study of Trends in Extreme Rainfall over the Iberian Peninsula. *J. Clim.* **2011**, *24*, 1089–1105. [[CrossRef](#)]
7. Acero, F.J.; Gallego, M.C.; García, J.A. Multi-day rainfall trends over the Iberian Peninsula. *Theor. Appl. Climatol.* **2011**, *108*, 411–423. [[CrossRef](#)]
8. Acero, F.J.; Parey, S.; Hoang, T.T.H.; Dacunha-Castelle, D.; García, J.A.; Gallego, M.C. Non-stationary future return levels for extreme rainfall over Extremadura (southwestern Iberian Peninsula). *Hydrol. Sci. J.* **2017**, *62*, 1394–1411. [[CrossRef](#)]
9. Wi, S.; Valdés, J.B.; Steinschneider, S.; Kim, T.W. Non-stationary frequency analysis of extreme precipitation in South Korea using peaks-over-threshold and annual maxima. *Stoch. Environ. Res. Risk Assess.* **2015**, *30*, 583–606. [[CrossRef](#)]
10. García, A.; Martín, J.; Naranjo, L.; Acero, F.J. A Bayesian hierarchical spatio-temporal model for extreme rainfall in Extremadura (Spain). *Hydrol. Sci. J.* **2018**, *63*, 878–894. [[CrossRef](#)]
11. Ramos, A.A. Extreme value theory and the solar cycle. *Astron. Astrophys.* **2007**, *472*, 293–298. [[CrossRef](#)]
12. Acero, F.J.; Carrasco, V.M.S.; Gallego, M.C.; García, J.A.; Vaquero, J.M. Extreme Value Theory and the New Sunspot Number Series. *Astrophys. J.* **2017**, *839*, 98. [[CrossRef](#)]
13. Acero, F.J.; Gallego, M.C.; García, J.A.; Usoskin, I.G.; Vaquero, J.M. Extreme Value Theory Applied to the Millennial Sunspot Number Series. *Astrophys. J.* **2018**, *853*, 80. [[CrossRef](#)]
14. Castillo, E.; Hadi, A.S.; Balakrishnan, N.; Sarabia, J.M. *Extreme Value and Related Models with Applications in Engineering and Science*; Wiley: Hoboken, NJ, USA, 2004.
15. Castillo, E. Estadística de valores extremos. Distribuciones asintóticas. *Estad. Esp.* **1987**, *116*, 5–35.
16. Smith, R.L.; Naylor, J.C. A Comparison of Maximum Likelihood and Bayesian Estimators for the Three-Parameter Weibull Distribution. *Appl. Stat.* **1987**, *36*, 358. [[CrossRef](#)]
17. Coles, S.; Pericchi, L.R.; Sisson, S. A fully probabilistic approach to extreme rainfall modeling. *J. Hydrol.* **2003**, *273*, 35–50. [[CrossRef](#)]
18. Bernardo, J.M.; Smith, A.F.M. (Eds.) *Bayesian Theory*; Wiley: Hoboken, NJ, USA, 1994. [[CrossRef](#)]
19. Kotz, S.; Nadarajah, S. *Extreme Value Distributions: Theory and Applications*; ICP: London, UK, 2000.
20. Coles, S.G.; Tawn, J.A. A Bayesian Analysis of Extreme Rainfall Data. *Appl. Stat.* **1996**, *45*, 463. [[CrossRef](#)]
21. Rostami, M.; Adam, M.B. Analyses of prior selections for gumbel distribution. *Matematika* **2013**, *29*, 95–107. [[CrossRef](#)]
22. Chen, M.H.; Shao, Q.M.; Ibrahim, J.G. *Monte Carlo Methods in Bayesian Computation*; Springer: New York, NY, USA, 2000. [[CrossRef](#)]
23. Vidal, I. A Bayesian analysis of the Gumbel distribution: An application to extreme rainfall data. *Stoch. Environ. Res. Risk Assess.* **2013**, *28*, 571–582. [[CrossRef](#)]
24. Lye, L.M. Bayes estimate of the probability of exceedance of annual floods. *Stoch. Hydrol. Hydraul.* **1990**, *4*, 55–64. [[CrossRef](#)]
25. Rostami, M.; Adam, M.B.; Ibrahim, N.A.; Yahya, M.H. Slice sampling technique in Bayesian extreme of gold price modelling. *Am. Inst. Phys.* **2013**, *1557*, 473–477. [[CrossRef](#)]
26. Gumbel, E.J. *Statistics of Extremes (Dover Books on Mathematics)*; Dover Publications: New York, NY, USA, 2012.
27. Gnedenko, B. Sur la distribution limite du terme maximum d'une serie aleatoire. *Ann. Math.* **1943**, *44*, 423–453. [[CrossRef](#)]

28. Fisher, R.A.; Tippett, L.H.C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*; Cambridge University Press: Cambridge, UK, 1928; Volume 24, pp. 180–190.
29. Ferreira, A.; de Haan, L. *Extreme Value Theory*; Springer: Berlin/Heidelberg, Germany, 2006.
30. Plummer, M.; Best, N.; Cowles, K.; Vines, K. CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News* **2006**, *6*, 7–11.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).