REVSTAT – Statistical Journal
Volume 0, Number 0, Month 0000, 000–000
https://doi.org/00.00000/revstat.v00i0.000

Selection of objective priors for Gumbel distribution parameters with application to maximum rainfall data

Authors: Fernando Antonio Moala
– Department of Statistics, State University of Sao Paulo, Brazil (f.moala@unesp.br)

Adriano Buran Moala
– Department of Statistics, State University of Sao Paulo, Brazil (adriano.moala@portoseguro.com.br)

Nixon Jerez-Lillo ✉
– Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Chile (njerez1@mat.uc.cl)

Pedro Luiz Ramos
– Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Chile (pedro.ramos@mat.uc.cl)

Received: Month 0000 Revised: Month 0000 Accepted: Month 0000

Abstract:

- The selection of priors is a critical aspect of Bayesian analysis, although the literature lacks studies concerning the application of the Gumbel distribution using different objective priors. We derive objective priors for the two-parameter Gumbel distribution and present a fully Bayesian analysis. Our primary goal is to choose a prior that represents a state of "little knowledge" a priori for both parameters. To yield this, we implement Markov Chain Monte Carlo algorithms to sample from the posterior distribution and to calculate the Bayes estimators. This investigation is made in the context of extreme weather events, using maximum rainfall data.

Keywords:

- *Gumbel distribution; Bayesian inference; objective priors; reference prior.*

AMS Subject Classification:

- 62F15, 62G32.

✉ Corresponding author

1. INTRODUCTION

The Gumbel distribution is widely used in various areas of knowledge to model the maximum (or minimum) values of occurrence for a given phenomenon of interest. Examples include extreme natural events in climatology (floods, earthquakes, climate changes), hydrology (river discharge volumes), geology, environmental science (concentration of pollution in the air), finance (study of financial crises), insurance (for large claims due to catastrophic events), resistance of materials, aeronautics, and naval engineering, among others.

In the field of climatology, numerous studies focus on examining extreme rainfall and predicting its future behavior utilizing the Gumbel distribution across various geographical zones, including South Korea (Nadarajah and Choi, 2007), Zimbabwe (Onwuegbuche et al., 2019), Kenya (Nadarajah and Choi, 2007), and Chile (Vidal, 2014). Analyzing the data distribution of maximum rainfall holds crucial importance for the development of agricultural and hydraulic engineering projects. This includes the design of irrigation and drainage channels, soil conservation efforts, roads, and dams. Accurate sizing necessitates a thorough understanding of extreme occurrences.

Let X be a random variable under the Gumbel distribution with the probability density function (pdf) given by

$$f(x \mid \mu, \varphi) = \frac{1}{\varphi} \exp\left(-\frac{x-\mu}{\varphi} - \exp\left(-\frac{x-\mu}{\varphi}\right)\right), \quad -\infty < x < \infty \tag{1.1}$$

with $-\infty < \mu < \infty$, $\varphi > 0$, where μ and φ are the location and scale parameters, respectively. The mean of this distribution is given by $E(X) = \mu + \gamma\varphi$, where γ is the Euler–Mascheroni constant ($\gamma \approx 0.5772$). The mode equals μ, the median is $\mu + \varphi \log(\log 2)$, and the variance is $\frac{\pi^2}{6}\varphi \approx 1.645\varphi$ (Kotz and Nadarajah, 2000).

Several methods can be used to obtain estimators for the parameters of the Gumbel distribution. The moments or maximum likelihood approaches are the most commonly used classical methods. Bayesian inference is another widely adopted approach, requiring the definition of a prior distribution for the parameters of the Gumbel distribution. This approach is especially useful when dealing with small sample sizes.

However, the practical use of Bayesian estimation often is associated with difficulties in choosing the prior distribution for the parameters. For example, Rostami and Adam (2013) compared eighteen pairs of independent proper priors, concluding that the combination of Gumbel and Rayleigh is the most productive pair of priors for the Gumbel model. Coles (2003) used a Bayesian approach to estimate Gumbel parameters by choosing normal and log-normal as priors for daily rainfall recorded data in Venezuela.

An important class of priors, known as objective priors (some of which can be seen as non-informative), is considered when we aim to construct a prior using a formal rule based on the model, rather than assessing it from expert opinions (see, for instance, Consonni et al., 2018; Kass and Wasserman, 1996). An objective prior allows the data, through the likelihood function, to fully dictate the posterior distribution, especially when little prior information about the parameters is available (Bernardo, 2005). Many of these prior distributions have been proposed by Jeffreys (1946), Bernardo (1979, 2005), Zellner (1977, 1984), Tibshirani

(1989), and others. The extensive number of objective priors, as well as their usage, has been derived for other models, as different priors may return better results for the posterior estimates in terms of bias, variance, and coverage probabilities when considering different models (see Tomazella et al., 2020; Ferreira et al., 2020; Ramos et al., 2021; Ramos et al., 2022).

Vidal (2014) has modeled annual rainfall maximum intensities by using noninformative prior distribution for the parameter μ and three different prior distributions for the parameter φ. Mousa et al. (2002) conducted a Bayesian method to examine the characterization, prediction and estimation of Gumbel model based on records data. Actually, they proposed a Jeffreys noninformative prior distribution for μ and a conjugate prior distribution for φ. Chbab et al. (2000) proposed the noninformative joint Jeffreys prior representing weak information for the parameters μ and φ. Lye et al. (1993) and Elkahlout (2006) also used noninformative Jeffreys prior to estimating the reliability function of the Gumbel distribution.

More recently, Yılmaz et al. (2021) compared many classical methods, such as the maximum likelihood, moment, least squares, weighted least squares, percentile, L-moments, trimmed L-moments estimators, best linear unbiased estimators, and Bain and Engelhardt estimators with the Bayesian method assuming weakly non-informative gamma priors and concluded that the Bayesian approach returned more accurate results. However, one of the priors is improper, and they did not prove that the resulting posterior would be proper, thus lacking theoretical justification. Meanwhile, the other priors were highly informative, which could lead to biased information in regions of the parameter space where the priors assigned small probabilities. Additionally, Bernardo (2005) argued that using simple proper priors as supposed to be non-informative, often hides significant unwarranted assumptions, which may easily dominate, or even invalidate the statistical analysis and should be strongly discouraged, and objective priors obtained by formal rules should be used instead.

In this paper, our focus is on the derivation and comparison of objective priors such as Jeffreys (Jeffreys, 1946), maximal data information (MDI) prior (Zellner, 1977), reference (Bernardo, 2005; Berger et al., 2015), and Tibshirani (Tibshirani, 1989) priors for the Gumbel distribution when both parameters μ and φ are of interest. It is worth mentioning that, although the Jeffreys priors and the reference priors are the most commonly presented in the statistical literature, the MDI and the priors proposed by Tibshirani, also known as matching priors, are also very useful for constructing objective priors. Comparing these priors, particularly for small sample sizes, is desirable. The motivation for this work partly lies in the importance of these priors in Bayesian inference, which can be used to obtain the parameter estimation objectively. It is also due to the necessity of considering both parameters unknown, i.e., without any information about them, not even the form of the prior distribution. We also demonstrated that these priors belong to a general class of power priors, in which the exponent determines the type of prior under consideration. These findings enabled us to identify conditions under which improper priors result in proper posteriors. More importantly, the obtained posterior returns better numerical results and also excellent theoretical properties such as invariance property under one-to-one transformations of the parameters, consistency under marginalization and consistent sampling properties.

As the marginal posterior densities and related Bayesian estimators cannot be obtained in closed form, we conduct Bayesian computations using Markov Chain Monte Carlo

(MCMC) techniques to generate samples from the joint posterior distributions. We then conduct a simulation study with various sample sizes to examine the bias, mean square error, and frequentist coverage probabilities, comparing the performance of the proposed priors. Additionally, we apply this approach to rainfall data to model maximum rainfall, which is crucial for planning in hydraulic engineering.

The remaining sections are organized as follows. Section 2 reviews the construction of the priors and their applications for the Gumbel parameters μ and φ. In Section 3, the joint and marginal posteriors for all proposed priors are also derived. The implementation of the MCMC algorithm applied to obtain the posterior summaries is given in section 4. Section 5 presents a numerical simulation to compare the performance of the prior distributions proposed in this paper. Section 6 illustrates the proposed Bayesian approach applied to rainfall data. Finally, Section 7 presents some conclusions and discussions.

2. OBJECTIVE PRIORS

There are a large number of objective priors that can be used in a Bayesian analysis, further we discuss the most commonly used prior.

2.1. Jeffreys prior

A well-known weak prior to representing a situation with little information available a priori about the parameters was proposed by Jeffreys (1946). Since then Jeffreys prior has played an important role in Bayesian inference. Tiao and Box (1973) give an explanation of the derivation of non-informative Jeffreys priors in terms of "data translated" likelihood. This prior is derived as

$$\pi_J(\boldsymbol{\theta}) \propto \sqrt{\det \mathcal{I}(\boldsymbol{\theta})}, \tag{2.1}$$

where $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^k$ is a vector of parameters, and $\mathcal{I}(\boldsymbol{\theta})$ is the Fisher information matrix with elements given by $\mathcal{I}_{ij}(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 \log f(x|\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right)$, for $i, j = 1, \ldots, k$. The Fisher information for the Gumbel distribution is given by

$$\mathcal{I}(\mu, \varphi) = \begin{bmatrix} \frac{1}{\varphi^2} & -\frac{1}{\varphi^2}[1+\psi(1)] \\ -\frac{1}{\varphi^2}[1+\psi(1)] & \frac{1}{\varphi^2}[1+\psi'(2)+\psi^2(2)] \end{bmatrix}, \tag{2.2}$$

where $\psi(k)$ is the digamma function, which is defined as the logarithmic derivative of the gamma function, i.e, $\psi(k) = \frac{d}{dk}\log\Gamma(k)$.

The rule of Jeffreys leads to the following joint prior distribution (see for instance Elkahlout, 2006) given by

$$\pi_J(\mu, \varphi) \propto \frac{1}{\varphi^2}. \tag{2.3}$$

2.2. Reference prior

Another well-known class of noninformative priors is the reference prior, first described by Bernardo (1979) and further developed by Berger et al. (1992). The idea is to derive a prior $\pi(\boldsymbol{\theta})$ that maximizes the expected posterior information about the parameters, provided by independent replications of an experiment relative to the information in the prior. A natural measure of the expected information about $\boldsymbol{\theta}$ provided by the data $\boldsymbol{x}$ is given by

$$I(\boldsymbol{\theta}) = E_x[K(p(\boldsymbol{\theta} \,|\, \boldsymbol{x}), \pi(\boldsymbol{\theta}))],$$

where

$$K(p(\boldsymbol{\theta} \mid \boldsymbol{x}),\, \pi(\boldsymbol{\theta})) = \int_{\Theta} p(\boldsymbol{\theta} \mid \boldsymbol{x}) \log\left(\frac{p(\boldsymbol{\theta} \mid \boldsymbol{x})}{\pi(\boldsymbol{\theta})}\right) d\boldsymbol{\theta}$$

is the Kullback-Leibler distance. Thus, the reference prior is defined as the prior $\pi(\boldsymbol{\theta})$ that maximizes the expected Kullback-Leibler distance between the posterior distribution and the prior distribution $\pi(\boldsymbol{\theta})$, taken over the experimental data.

The method leads to Jeffreys prior in the one-dimensional case, but it has some advantages over Jeffreys prior in the multidimensional case (see Berger et al., 1992 for details).

The general reference prior is typically very hard to implement; however, for the regular case, in which the asymptotic normality of the model holds, a considerable simplification of the algorithm occurs. Zhang and Shaby (2023) provided a comprehensive analysis of the asymptotic normality of the Generalized Gumbel distribution, which includes the Gumbel distribution as a subclass. They demonstrated the asymptotic posterior normality of the model, a pivotal result elaborated in Proposition 3.1 of their study.

An important feature in this approach is the different treatment for interest and nuisance parameters when $\boldsymbol{\theta}$ is a vector of parameters. In the presence of nuisance parameters, one must establish an ordered parameterization with the parameter of interest singled out and then follow the procedure below. The algorithm of Berger et al. (1992) to derive the reference prior, in the two-parameter case, can be described in four steps, as follows.

Let $\boldsymbol{\theta} = (\theta_1, \theta_2)$ be the whole parameter, with θ_1 being the parameter of interest and θ_2 the nuisance parameter. The algorithm is as follows:

Step 1: Determine $\pi_2(\theta_2 \mid \theta_1)$, the conditional reference prior for θ_2 assuming that θ_1 is given, i.e.,

$$\pi_2(\theta_2 |\, \theta_1) = \sqrt{\mathcal{I}_{22}(\theta_1, \theta_2)}$$

where $\mathcal{I}_{22}(\theta_1, \theta_2)$ is the (2,2)-entry of the Fisher Information Matrix.

Step 2: Normalize $\pi_2(\theta_2 \mid \theta_1)$. In case that $\pi_2(\theta_2 \mid \theta_1)$ was improper, choose a sequence of subsets $\Theta_1 \subseteq \Theta_2 \subseteq \ldots \to \Theta$ on which $\pi_2(\theta_2 \mid \theta_1)$ is proper. Then, define

$$c_m(\theta_1) = \frac{1}{\int_{\Theta_m} \pi_2(\theta_2 |\, \theta_1) d\theta_2},$$

and

$$p_m(\theta_2 |\, \theta_1) = c_m(\theta_1) \pi_2(\theta_2 |\, \theta_1) \mathbb{I}_{\Theta_m}(\theta_2).$$

Step 3: Find the marginal reference prior for θ_1, i.e., the reference prior for the experiment formed by marginalizing out with respect to $p_m(\theta_2|\,\theta_1)$. We obtain

$$\pi_m(\theta_1) \propto \exp\left\{\frac{1}{2}\int_{\Theta_m} p_m(\theta_2|\,\theta_1)\log\left\|\frac{\det\mathcal{I}(\theta_1,\theta_2)}{\mathcal{I}_{22}(\theta_1,\theta_2)}\right\|d\theta_2\right\}. \tag{2.4}$$

Step 4: Compute the reference prior for (θ_1, θ_2) when θ_2 is a nuisance parameter:

$$\pi_B(\theta_1,\theta_2) = \lim_{m\to\infty}\left(\frac{c_m(\theta_1)\pi_m(\theta_1)}{c_m(\theta_1^*)\pi_m(\theta_1^*)}\right)\pi(\theta_2 \mid \theta_1)$$

where θ_1^* is any fixed point with positive density for all π_m;

Theorem 2.1. *The reference prior for the parameters μ and φ of the Gumbel distribution (1.1) is given by*

$$\pi_B(\mu,\varphi) \propto \frac{1}{\varphi}. \tag{2.5}$$

Proof: The reference prior is obtained firstly for the nuisance parameter φ, conditionally on the parameter of interest μ, as

$$\pi(\varphi \mid \mu) \propto \sqrt{\mathcal{I}_{\varphi\varphi}(\mu,\,\varphi)} = \frac{1}{\varphi}$$

which is an improper prior.

A natural sequence of compact sets for (μ, φ) is $\Theta_m = (l_{1m}, l_{2m}) \times (q_{1m}, q_{2m})$, so that $l_{1m} \to -\infty$, $l_{1m} \to \infty$, $q_{1m} \to 0$, and $q_{2m} \to \infty$ when $m \to \infty$. Hence, the normalizing constant of prior $\pi(\varphi \mid \mu)$ is

$$c_m(\mu) = \frac{1}{\int_{q_{1m}}^{q_{2m}}\frac{1}{\varphi}d\varphi} = \frac{1}{\log q_{2m} - \log q_{1m}},$$

and $p_m(\varphi \mid \mu) = \frac{1}{\log q_{2m}-\log q_{1m}}\frac{1}{\varphi}1_{\Theta_m}(\varphi)$.

Now, from (2.4) the marginal reference prior for μ after some algebras is given by

$$\pi_m(\mu) \propto \exp\left(\int_{q_{1m}}^{q_{2m}}\log\left(\frac{1}{\varphi}\sqrt{\frac{1+\psi(2)}{1+\psi'(2)+\psi^2(2)}}\right)c_m(\mu)\frac{1}{\varphi}d\varphi\right)$$

that does not depend of μ.

Therefore the global referencial prior for $(\mu,\,\varphi)$ with parameter of interesting μ is given by,

$$\pi_\mu(\mu,\varphi) = \lim_{m\to\infty}\left(\frac{c_m(\mu)\pi_m(\mu)}{c_m(\mu^*)\pi_m(\mu^*)}\right)\pi(\varphi \mid \mu) = \frac{1}{\varphi},$$

considering $\mu^* = 1$.

In a similar way, we obtain the reference prior considering parameter of φ interest and μ nuisance. In this case, we define

$$\pi(\mu \mid \varphi) \propto \sqrt{\mathcal{I}_{\mu\mu}(\mu,\, \varphi)} \propto 1.$$

From the mentioned sequence of compact sets, it follows that the normalizing constant of prior $\pi(\mu \mid \varphi)$ is

$$d_m(\varphi) = \frac{1}{\int_{l_{1m}}^{l_{2m}} d\mu} = \frac{1}{l_{2m} - l_{1m}},$$

and $p_m(\mu \mid \varphi) = \frac{1}{l_{2m}-l_{1m}} 1_{\Theta_m}(\mu)$.

From (2.4) the marginal reference prior for φ is given by

$$\pi_m(\varphi) \propto \exp\left(\int_{l_{1m}}^{l_{2m}} \log\left(\frac{1}{\varphi}\sqrt{1+\psi'(2)}\right) d_m(\varphi) d\mu\right) \propto \frac{1}{\varphi}.$$

Therefore the global referencial prior for $(\mu,\, \varphi)$ with parameter of interesting φ is given by,

$$\pi_\varphi(\mu, \varphi) = \lim_{m\to\infty}\left(\frac{d_m(\varphi)\pi_m(\varphi)}{d_m(\varphi^*)\pi_m(\varphi^*)}\right)\pi(\mu \mid \varphi) \propto \frac{1}{\varphi},$$

considering $\varphi^* = 1$. This way, the reference prior with parameter of interesting φ is the same with μ as the parameter of interest, and hence,

$$\pi_B(\mu, \varphi) \propto \frac{1}{\varphi}.$$

□

Due to this fact, the reference prior presented above is an overall reference prior (see Berger et al., 2015), and all the parameters are of interest.

2.3. Maximal data information prior

Zellner (1977) developed a prior $\pi(\boldsymbol{\theta})$ that maximizes the information from the data in relation to the information known a priori about the parameters. Let

$$H(\boldsymbol{\theta}) = \int_{-\infty}^{\infty} f(x \mid \boldsymbol{\theta}) \log f(x \mid \boldsymbol{\theta}) dx$$

be the negative entropy of $f(x \mid \boldsymbol{\theta})$. The non-informative prior distribution known as Maximal Data Information prior is defined as

$$\pi_Z(\boldsymbol{\theta}) \propto \exp H(\boldsymbol{\theta}).$$

Zellner (1977, 1984) shows several interesting properties of MDI prior and additional conditions that can also be imposed on the approach reflecting given initial information. However, the MDI prior has restrictive invariance properties.

Theorem 2.2. *The MDI prior for the parameters μ and φ from Gumbel distribution* (1.1) *is given by:*

$$\pi_Z(\mu,\,\varphi) \propto \frac{1}{\varphi} \tag{2.6}$$

and, therefore, the MDI prior coincides with the reference prior.

Proof: Firstly, we have to evaluate the measure $H(\mu,\,\varphi)$. After some algebras,

$$\begin{aligned} H(\mu,\,\varphi) &= \int_{-\infty}^{\infty} f(x\mid\mu,\,\varphi)\log f(x\mid\mu,\,\varphi)\,dx, \\ &= \int_{-\infty}^{\infty} f(x\mid\mu,\,\varphi)\left[-\log\varphi - \frac{x-\mu}{\varphi} - \exp\left(-\frac{x-\mu}{\varphi}\right)\right]dx, \\ &= -\log\varphi - \frac{E(X)-\mu}{\varphi} - \int_{-\infty}^{\infty}\exp\left(-\frac{x-\mu}{\varphi}\right) f(x\mid\mu,\,\varphi)\,dx, \\ &= -\log\varphi - \gamma - 1. \end{aligned}$$

This way, the noninformative MDI prior for the parameters μ and φ is given by

$$\pi_Z(\mu,\,\varphi) \propto \frac{1}{\varphi}.$$

□

2.4. Tibishirani's prior

Tibshirani (1989) proposes a procedure for obtaining an objective prior for a parametric vector $\boldsymbol{\theta} = (\theta_1,\ldots,\,\theta_k)$ such that, for a fixed coefficient, the marginal posterior interval for a single parameter of interest has coverage probability in the frequentist sense close to the nominal significance level. Consider the vector of parameters $\boldsymbol{\theta} = (\theta_1, \boldsymbol{\theta}_2)$ in which the scalar θ_1 is the parameter of interest and $\boldsymbol{\theta}_2 = (\theta_{21}, \theta_{22}, \ldots, \theta_{2,k-1})$ is the vector of nuisance parameters. Here we want to obtain, a non-informative prior $\pi(\theta_1, \theta_2)$ where θ_1 is the parameter of interest so that the credible interval for θ_1 has a coverage error $O(n^{-1})$ in the frequentist sense, i.e.,

$$P\left[\theta_1 \le \theta_1^{1-\alpha}(\pi;X)|(\theta_1,\theta_2)\right] = 1-\alpha-O(n^{-1}), \tag{2.7}$$

where $\theta_1^{1-\alpha}(\pi;X)|(\theta_1,\theta_2)$ denote the $(1-\alpha)$th quantile of the posterior distribution of θ_1. The class of priors satisfying (2.7) are known as matching priors.

The procedure requires an orthogonal reparametrization of the parameter $\boldsymbol{\theta}$. To achieve this, we will employ the criterion proposed by Cox and Reid (1987), involving the reparametrization of the model in a system with $k-1$ equations, as follows:

$$\sum_{i=1}^{k-1} i^*_{\theta_{2r}\theta_{2s}} \frac{\partial\theta_{2r}}{\partial\theta_1} = -i^*_{\theta_1\theta_{2s}}, \quad s = 1,\ldots,k-1, \tag{2.8}$$

where i^* are the elements of the information matrix obtained in the parametrization $(\theta_1, \boldsymbol{\theta}_2)$ under the assumption that the transformation from $(\theta_1, \boldsymbol{\theta}_2)\to(\delta, \boldsymbol{\lambda})$ has a non-null Jacobiano and is orthogonal. Therefore, the $k \times k$ Fisher information matrix for the new orthogonal parameters δ and $\boldsymbol{\lambda}$, which corresponds to the parameter of interest δ is given by

$$\mathcal{I}(\delta, \boldsymbol{\lambda}) = \begin{bmatrix} \mathcal{I}_{\delta\delta}(\delta, \boldsymbol{\lambda}) & 0 \\ 0 & \mathcal{I}_{\boldsymbol{\lambda}\boldsymbol{\lambda}}(\delta, \boldsymbol{\lambda}) \end{bmatrix},$$

where

$$\mathcal{I}_{\delta\delta}(\delta, \boldsymbol{\lambda}) = -E\left(\frac{\partial^2 L(\delta, \boldsymbol{\lambda} \mid \boldsymbol{x})}{\partial \delta^2}\right), \quad [\mathcal{I}_{\boldsymbol{\lambda}\boldsymbol{\lambda}}(\delta, \boldsymbol{\lambda})]_{ij} = -E\left(\frac{\partial^2 L(\delta, \boldsymbol{\lambda} \mid \boldsymbol{x})}{\partial \lambda_i \partial \lambda_j}\right).$$

This way, the prior given by

$$\pi(\delta, \boldsymbol{\lambda}) \propto g(\boldsymbol{\lambda}) \sqrt{\mathcal{I}_{\delta\delta}(\delta, \boldsymbol{\lambda})}, \tag{2.9}$$

leads (asymptotically in the sample size) to the proper frequentist coverage one-sided posterior interval for δ and $g(\boldsymbol{\lambda}) > 0$ is an arbitrary function. The corresponding prior for $\boldsymbol{\theta}$ is given by

$$\pi_T(\boldsymbol{\theta}) \propto g(\boldsymbol{\lambda}) \mathcal{I}_{\delta\delta}(\delta, \boldsymbol{\lambda})^{1/2} \left\| J(\delta, \boldsymbol{\lambda}) \right\|,$$

where $\|J(\delta, \boldsymbol{\lambda})\|$ is the Jacobian of the transformation from $(\delta, \boldsymbol{\lambda})$ to $\boldsymbol{\theta}$.

Due to a lack of uniqueness in the choice of the orthogonal parametrization then the class of orthogonal parameters is of the form $g(\lambda)$, where $g(\cdot)$ is any reparametrization. This nonuniqueness is reflected by the function $g(\cdot)$ corresponding to (2.9). One possibility, in the single nuisance parameter case, is to require that (δ, λ) also satisfies Stein's condition (see Tibshirani, 1989) for λ with δ taken as the nuisance parameter.

Theorem 2.3. *The Tibshirani prior for the parameters μ and φ of the extreme value distribution* (1.1) *is given by*

$$\pi_T(\mu, \varphi) \propto \frac{1}{\varphi}. \tag{2.10}$$

Proof: As discussed by Cox and Reid (1987), the orthogonal parametrization diagonalizes the Fisher matrix (2.2). Therefore, we transform the original parameters (μ, φ) to the orthogonal parameters (δ, λ), by solving the differential equation in (2.8), which in our case is given by

$$\frac{1}{\varphi^2} \frac{\partial \mu}{\partial \varphi} = \frac{a}{\varphi^2},$$

where $a = 1 + \psi(1)$ is a constant.

The solution of the differential equation gives an one-to-one transformation

$$\lambda = \mu - a\varphi. \tag{2.11}$$

From (2.2) and (2.11), the Fisher Information matrix for the orthogonal parameters is given by

$$I(\delta, \lambda) = \frac{1}{\delta^2} \begin{bmatrix} b & 0 \\ 0 & 1 \end{bmatrix},$$

where $b = -a^2 + \big(1 + \psi'(2) + \psi^2(2)\big)$.

From (2.9) and (2.11), the corresponding prior for (δ, λ) is given by

$$\pi(\delta, \lambda) \propto g(\lambda)\frac{1}{\delta},$$

where $g(\lambda)$ is an arbitrary function.

Now, assuming the parameter δ as the single nuisance parameter we obtain

$$\pi(\delta,\, \lambda) \propto g^*(\delta)\frac{1}{\delta}.$$

Now, requiring $g(\lambda)\frac{1}{\delta} = g^*(\delta)\frac{1}{\delta}$ we have that

$$\pi(\delta,\, \lambda) \propto \frac{1}{\delta}.$$

Thus, the prior expressed in terms of the $(\mu,\, \varphi)$ parametrization is given by

$$\pi_T(\mu,\, \varphi) \propto \frac{1}{\varphi}.$$

□

3. POSTERIOR DISTRIBUTION

Suppose that an objective Bayesian approach will be considered, the priors discussed so far can be represented as a class of priors given by

$$\pi(\mu,\, \varphi) \propto \frac{1}{\varphi^k}, \tag{3.1}$$

where k is a known constant that depends on the chosen prior, i.e, the uniform prior ($k = 0$), reference, MDI and Tibishirani ($k = 1$), and finally Jeffreys ($k = 2$).

Note that the priors (3.1) are particular cases of joint independent prior given by $\pi(\mu,\, \varphi) = \pi(\mu)\pi(\varphi)$ with $\pi(\mu) \propto 1$ and $\pi(\varphi) \propto \frac{1}{\varphi^{a+1}}\exp(-\frac{b}{\varphi})$.

By considering a random sample $x_1, \ldots, x_n$ from the Gumbel distribution with pdf given in (1.1), then the likelihood function for the parameters μ and φ, based on $\boldsymbol{x}$, is given by

$$L(\mu,\, \varphi \mid \boldsymbol{x}) \;=\; \frac{1}{\varphi^n}\exp\left(-\frac{n}{\varphi}(\overline{x} - \mu) - \sum_{i=1}^{n}\exp\left(-\frac{x_i - \mu}{\varphi}\right)\right), \tag{3.2}$$

where $\overline{x} = \sum_{i=1}^{n} \frac{x_i}{n}$ is the sample mean. Now, from (3.2) and (3.1), the joint posterior distribution for $(\mu,\, \varphi)$ is given by

$$p(\mu,\, \varphi \mid \boldsymbol{x}) \propto \frac{1}{\varphi^{n+k}}\exp\left(-\frac{n}{\varphi}(\overline{x} - \mu) - \sum_{i=1}^{n}\exp\left(-\frac{x_i - \mu}{\varphi}\right)\right). \tag{3.3}$$

As the posterior could return an improper posterior, the following proposition shows us that under the assumptions of $n+k>2$ any of the priors returns proper posterior.

Theorem: The joint posterior density $p(\mu,\, \varphi \mid \boldsymbol{x})$ for parameters $(\mu,\, \varphi)$ is proper if and only if $n+k>2$ for any $k=0,\, 1,\, 2$.

Proof: We are interested in proving that the double integral of the joint posterior (3.3) is finite, i.e,

$$\int_0^{\infty}\int_{-\infty}^{+\infty} p(\mu,\, \varphi \mid \boldsymbol{x}) d\mu d\varphi < \infty.$$

Integrating first with respect to μ, we have

(3.4)
$$\int_{-\infty}^{+\infty} p(\mu,\, \varphi \mid \boldsymbol{x}) d\mu = \frac{1}{\varphi^{n+k}} \exp\left(-\frac{n\overline{x}}{\varphi}\right) \int_{-\infty}^{+\infty} \exp\left\{\frac{n\mu}{\varphi} - \exp\left(\frac{\mu}{\varphi}\right)\sum_{i=1}^{n}\exp\left(-\frac{x_i}{\varphi}\right)\right\} d\mu.$$

To compute the integral in (3.4), we will consider the transformation given by $u = \exp\left(\frac{\mu}{\varphi}\right)$ as a function of μ, then we obtain that

$$\int_{-\infty}^{+\infty} \exp\left\{\frac{n\mu}{\varphi} - \exp\left(\frac{\mu}{\varphi}\right)\Lambda\right\} d\mu = \varphi\int_0^{\infty} u^{n-1}\exp(-\Lambda u) du,$$

which is a kernel of a gamma distribution parameterized in terms of a shape parameter n and an rate parameter $\Lambda = \sum_{i=1}^{n}\exp\left(-\frac{x_i}{\varphi}\right)$. Therefore

$$\int_{-\infty}^{+\infty} p(\mu,\, \varphi \mid \boldsymbol{x}) d\mu = \frac{1}{\varphi^{n+k-1}}\exp\left(-\frac{n\overline{x}}{\varphi}\right)\frac{\Gamma(n)}{\Lambda^n}.$$

The double integral of joint posterior density for parameters μ and φ can be written as

$$\int_0^{\infty}\int_{-\infty}^{+\infty} p(\mu,\, \varphi \mid \boldsymbol{x}) d\mu d\varphi = \Gamma(n)\int_0^{\infty} \frac{\exp\left(-\frac{n\overline{x}}{\varphi}\right)}{\varphi^{n+k-1}\left(\sum_{i=1}^{n}\exp\left(-\frac{x_i}{\varphi}\right)\right)^n} d\varphi. \tag{3.5}$$

This way, we just need to prove that integral in (3.5) is finite. In fact, consider a new transformation $w = \frac{1}{\varphi}$ in the integral (3.5), this way we have

$$\int_0^{\infty} \frac{\exp\left(-\frac{n\overline{x}}{\varphi}\right)}{\varphi^{n+k-1}\left(\sum_{i=1}^{n}\exp\left(-\frac{x_i}{\varphi}\right)\right)^n} d\varphi = \int_0^{\infty} w^{n+k-3}\frac{\exp\left(-wn\overline{x}\right)}{\left(\sum_{i=1}^{n}\exp(-wx_i)\right)^n} dw.$$

Let $x_{(1)}$ be the sample minimum. It is clear that $\sum_{i=1}^{n}\exp(-wx_i) \geq \exp(-wx_{(1)})$ and $\left[\sum_{i=1}^{n}\exp(-wx_i)\right]^{-1} \leq \left[\exp(-wx_{(1)})\right]^{-1}$. Therefore

$$\int_0^{\infty} w^{n+k-3}\frac{\exp\left(-wn\overline{x}\right)}{\left(\sum_{i=1}^{n}\exp(-wx_i)\right)^n} dw \leq \int_0^{\infty} w^{(n+k-2)-1}\exp\Big(-wn(\overline{x}-x_{(1)})\Big) dw. \tag{3.6}$$

Hence, we have that (3.6) is finite if and only if $\overline{x} > x_{(1)}$ and $n+k > 2$. Consequently, the joint posterior is proper if $n > 2$ under the uniform prior ($k = 0$), if $n > 1$ under the reference, MDI and Tibishirani prior ($k = 1$), and for all n under the Jeffreys prior ($k = 2$). □

A corollary of this demonstration relates to the existence of the moments of φ. The p-th moment of φ is calculated as

$$E(\varphi^p \mid \boldsymbol{x}) \propto \int_0^{\infty} \frac{\exp\left(-\frac{n\overline{x}}{\varphi}\right)}{\varphi^{n+k-p-1}\left(\sum_{i=1}^{n}\exp\left(-\frac{x_i}{\varphi}\right)\right)^n} d\varphi.$$

Therefore, $E(\varphi^p \mid \boldsymbol{x}) < \infty$ if $n + k - p > 2$. For instance, the expectation and the variance exist if $n + k > 3$ and $n + k > 4$, respectively.

Now the inferences about μ and φ are typically based on marginal posterior distributions.

Proposition 4: The marginal posterior distribution for the parameter φ is

$$p(\varphi \mid \boldsymbol{x}) \propto \frac{\exp\left(-\frac{n\overline{x}}{\varphi}\right)}{\varphi^{n+k-1}\left(\sum_{i=1}^{n}\exp\left(-\frac{x_i}{\varphi}\right)\right)^n}. \tag{3.7}$$

Proof: For definition and from the joint posterior (3.3), the marginal posterior distribution for the parameter φ is obtained as

$$p(\varphi \mid \boldsymbol{x}) \propto \frac{1}{\varphi^{n+k}} \exp\left(-\frac{n\overline{x}}{\varphi}\right) \int_{-\infty}^{\infty} \exp\left(\frac{n\mu}{\varphi} - \sum_{i=1}^{n}\exp\left(-\frac{x_i-\mu}{\varphi}\right)\right) d\mu.$$

Making the change of variable $u = \frac{\mu}{\varphi}$ and some algebraic manipulations, the marginal distribution for φ is obtained in a closed form given by

$$p(\varphi \mid \boldsymbol{x}) \propto \frac{\exp\left(-\frac{n\overline{x}}{\varphi}\right)}{\varphi^{n+k-1}\left(\sum_{i=1}^{n}\exp\left(-\frac{x_i}{\varphi}\right)\right)^n}.$$

The proof is completed. □

On the other hand, the marginal posterior for μ is also derived from joint posterior (3.3) by direct integration out φ as

$$p(\mu \mid \boldsymbol{x}) \propto \int_0^{\infty} \frac{1}{\varphi^{n+k}} \exp\left(-\frac{n}{\varphi}(\overline{x}-\mu) - \sum_{i=1}^{n}\exp\left(-\frac{x_i-\mu}{\varphi}\right)\right) d\varphi. \tag{3.8}$$

Unfortunately, it is intractable to compute the normalization constant, Bayes summaries and intervals. However, samples of the marginal posterior distributions (3.7) and (3.8) can be simulated by using MCMC (Markov Chain Monte Carlo) methods, as the popular the Metropolis-Hastings algorithm (see for example, Chib and Greenberg, 1995).

4. IMPLEMENTING THE MCMC ALGORITHM

Since the joint posterior distribution in (3.3) do not have closed form, MCMC algorithms, in particular the Metropolis-Hasting (MH), can be implemented in this case quite routinely for the computation of Bayes estimators and quantiles.

Applying the Bayes theorem, we can write the joint posterior distribution as the product between the marginal distribution of $(\varphi \mid \boldsymbol{x})$ in (3.7) and the conditional distribution $(\mu \mid \varphi, \boldsymbol{x})$, i.e,

$$p(\mu, \varphi \mid \boldsymbol{x}) = p(\mu \mid \varphi, \boldsymbol{x}) p(\varphi \mid \boldsymbol{x}).$$

As none of the previous distributions are identified as having a standard shape, disabling the direct generation of samples, the MH algorithm will be considered to sample both conditional and marginal distributions. We choose the Normal $N(\mu_0, \sigma^2)$ and Gamma $G(a, b)$ distributions as the proposal distributions for the chain of μ and φ, respectively. The selection was only for the fact that for the μ we have the values on $(-\infty, \infty)$ while φ is defined on $(0, \infty)$, however, other distributions could be assumed respecting the parameter space.

The algorithm used to obtain the Bayesian estimates consists of the following steps

Step 1: Take some initial guess μ_o and φ_o for the parameters μ and φ, respectively;

Step 2: At the step i, we draw the new values μ_{i+1} and φ_{i+1} from $N(\mu_i, \sigma^2)$ and $G(b/\varphi_i{+}1, b)$ distributions, respectively, where σ and b, are scale parameters to balance the acceptance rate to about 0.45 according to Neal and Roberts (2008);

Step 3: The candidate φ_{i+1} will be accepted with a probability given by the Metropolis ratio

$$\alpha_2(\varphi_i, \varphi_{i+1}) = \min\left\{1, \frac{G(b/\varphi_i + 1, b)\, p\,(\varphi_{i+1} \mid \boldsymbol{x})}{G(b/\varphi_{i+1} + 1, b)\, p\,(\varphi_i \mid \boldsymbol{x}\,)}\right\},$$

and then, the candidate μ_{i+1} with probability

$$\alpha_1(\mu_i\,, \mu_{i+1}) = \min\left\{1, \frac{p\,(\mu_{i+1} \mid \varphi_{i+1}, \boldsymbol{x}\,)}{p\,(\mu_i \mid \varphi_{i+1}, \boldsymbol{x}\,)} = \frac{p\,(\mu_{i+1}, \varphi_{i+1} \mid \boldsymbol{x}\,)}{p\,(\mu_i, \varphi_{i+1} \mid \boldsymbol{x}\,)}\right\};$$

Step 4: Repeat above step N times to obtain one chain of N values for each one of μ and φ parameters;

Step 5: The Bayes estimators of the parameters μ and φ, under the square error loss function (SELF) are given by

$$\widehat{\mu} = \frac{1}{N-M} \sum_{i=M+1}^{N} \mu_i, \quad \text{and} \quad \widehat{\varphi} = \frac{1}{N-M} \sum_{i=M+1}^{N} \varphi_i,$$

respectively, where M is the burn-in period;

Step 6: Calculate the $100(1-\alpha)\%$ Bayesian interval $(Q_{(\frac{\alpha}{2})}, Q_{(1-\frac{\alpha}{2})})$, where $Q_{(k)}$ is the k-th quantile of μ and φ.

5. SIMULATION STUDY

In this section, we present a simulation study for the purpose of verifying the performance of the Bayesian estimators using the priors proposed in the previous sections. The R Core Team (2020) was used to develop the codes, which are accessible in a GitHub repository provided at the end of this paper. No special packages beyond those typically included in R were employed.

The simulation study was carried out for different sample sizes and with different parameter values of μ and φ. Specifically, we considered sample sizes of $n = 5, 10, 15, \ldots, 60$. We generated $N = 1,000$ samples from (1.1) for three arbitrarily chosen sets of parameter values, μ and φ, specifically $(\mu, \varphi) = (3, 1)$, $(\mu, \varphi) = (1, 2)$, and $(\mu, \varphi) = (0.5, 1)$.

We ran the MCMC algorithm described in Section 4 to obtain 5,500 observations of posterior distributions. The first 500 observations were discarded as the burn-in period, with a thinning interval of 5 draws. Therefore, we obtained two chains with the size of 1000 to compute the posterior summaries. The choice of initial values was made through the moment estimators given by Yılmaz et al. (2021):

$$\widehat{\mu}_0 = \overline{x} - \frac{\gamma s\sqrt{6}}{\pi}, \quad \widehat{\varphi}_0 = \frac{s\sqrt{6}}{\pi},$$

where $\overline{x}$ is the mean, s is the empirical variance and $\gamma \approx 0.5772$.

The convergence of the generated Markov chains was checked using the Geweke test (Geweke, 1992). This test is based on comparing the means of the first 10% and last 50% of the chain. It indicates that if the samples are drawn from the stationary distribution (the posterior distribution in our context), then the means of these two parts should be roughly equal. Geweke's statistic is calculated under the assumption that the two parts of the chain are asymptotically independent. Under this condition, the statistic follows an asymptotically standard normal distribution. Consequently, the chains are considered to have converged if the absolute value of the statistic is less than 1.96.

Additionally, we have included the results of the maximum likelihood estimators (MLEs) and the posterior distribution obtained using a highly informative prior considered in Yılmaz et al. (2021) to compare our results. This prior specifies a standard normal distribution for μ and a gamma distribution for φ with a mean equal to 1 and a variance equal to 0.25.

The MLEs were computed by maximizing the log-likelihood function (Elkahlout, 2006), which is derived by applying the logarithm to the likelihood function in (3.2). This approach follows the frequentist method of statistical inference (Bickel and Doksum, 2015). We report the empirical bias and the variance for each parameter

$$\text{Bias}(\theta_i) = \frac{1}{N}\sum_{j=1}^{N}(\hat{\theta}_{i,j} - \theta_i) \ , \quad \text{Var}(\theta_i) = \frac{1}{N}\sum_{j=1}^{N}(\hat{\theta}_{i,j} - \overline{\theta}_i)^2, \ \text{ for } \ i = 1, 2,$$

where $\overline{\theta}_i = \frac{1}{N}\sum_{j=1}^{N}\hat{\theta}_{i,j}$, and $\boldsymbol{\theta} = (\theta_1, \theta_2) = (\mu, \varphi)$. In addition, the 95% credibility interval coverage probabilities (CP) were computed for θ_i's, defined as the proportion of the total intervals covering the true parameter. Adequate performance would see these probabilities fall within the neighborhood of 0.95. The empirical bias, variance, and CP for all estimators versus sample sizes are plotted in Figures 1 to 3, respectively, for both different parameter values.

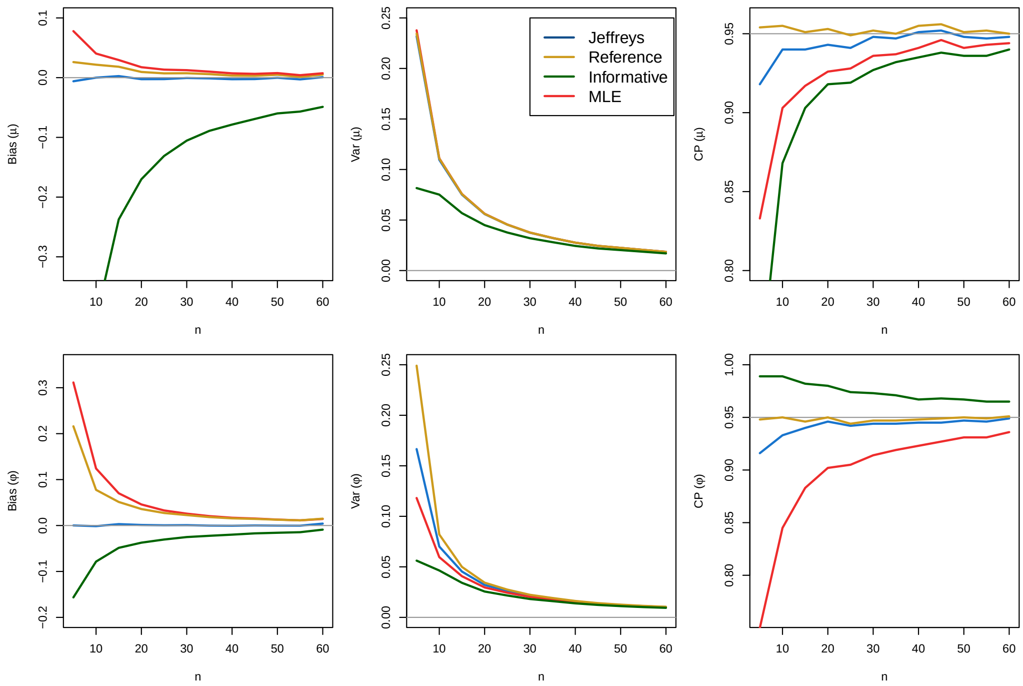

Figure 1: Plot of empirical bias, variance and coverage probabilities versus sample size n for estimation of $\mu = 3$ and $\varphi = 1$.

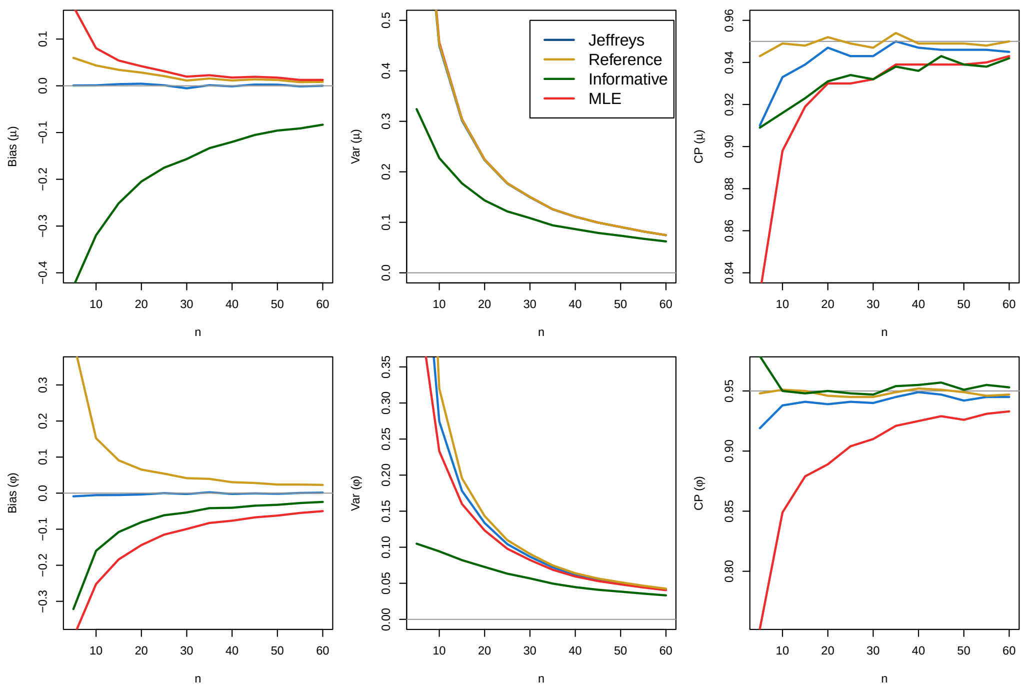

Figure 2: Plot of empirical bias, variance and coverage probabilities versus sample size n for estimation of $\mu = 1$ and $\varphi = 2$

Upon examining Figures 1 to 3, we conclude that some estimators, such as the MLE

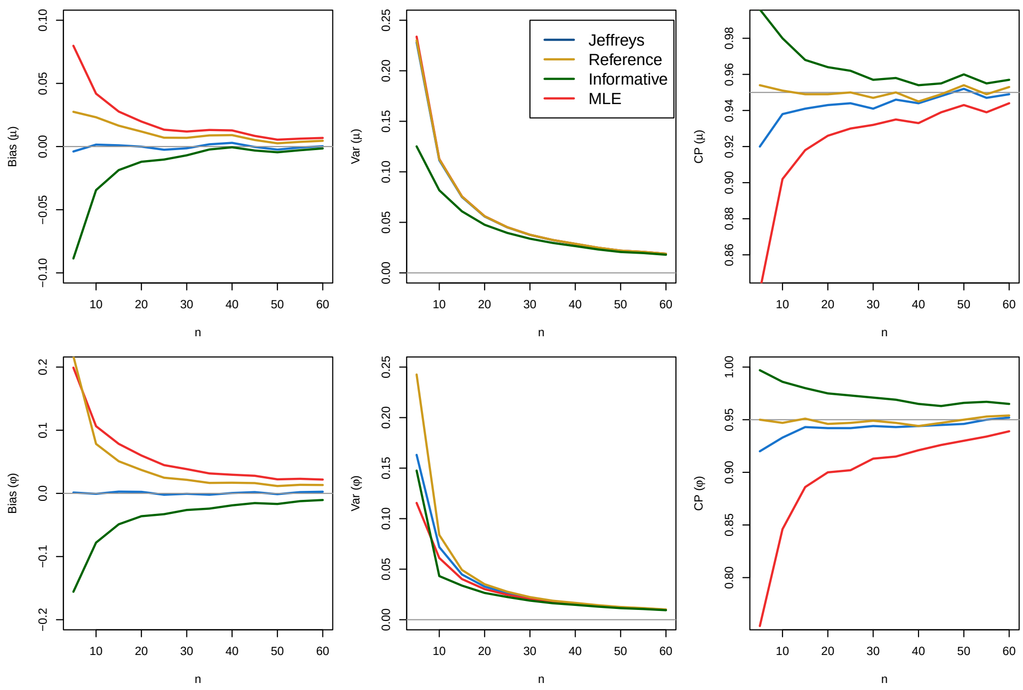

Figure 3: Plot of empirical bias, variance and coverage probabilities versus sample size n for estimation of $\mu = 0.5$ and $\varphi = 1$

and two priors (reference and informative), exhibit bias for small n for both parameter, while the Jeffreys prior returns nearly unbiased estimates. Most importantly, the bias for the informative prior requires a large n to decrease, while the biases of the objective priors tend towards zero for small samples. Additionally, the variance decreases as the sample size increases. Although the informative prior has the smallest variance, this is expected as the variance of the prior is small, providing less variability in the obtained estimates. This consequently leads to highly biased estimates, especially when they are far from the mean of the prior distribution. These results show that misspecifying the hyperparameters may lead to biased results, while our proposed Bayesian approach does not require any hyperparameters.

The effect of the proposed priors on the posterior estimates is also compared in terms of coverage probabilities (CP). As expected, as the sample size increases, the objective Bayesian credible intervals converge quickly to the nominal level of 0.95. On the other hand, informative priors tend to overestimate the nominal levels in almost all cases, resulting in poor coverage probabilities. The objective Bayesian estimator based on the reference prior performed better than the one based on the Jeffreys prior under this criterion, although both converged rapidly with relatively small sample sizes. Overall, we obtained unbiased estimates with accurate coverage probabilities when considering the objective Jeffreys prior, which is therefore recommended for future research in real applications.

It is worth mentioning that the Geweke test results are displayed in Figures 5 to 7 in the Appendix. From the cited figures, we conclude that all chains generated for any parameter across different scenarios and sample sizes exhibit a statistic with an absolute value less than 1.96, indicating convergence of the chains and validating our results.

6. APPLICATION TO RAINFALL DATA

Intense rainfall leads to flooding, road damage, soil erosion, and agricultural damage, among other effects. Understanding the behavior of maximum rainfall is crucial for planning in hydraulic engineering, such as flood control, dam construction, and agronomy, including determining the optimal time for planting and soil conservation. Plans for any such constructions must be based on extreme rainfall amounts. Therefore, given the socio-economic significance of the Presidente Prudente region, this application holds great relevance. Maximum rainfall data were obtained from the Meteorological Station of Presidente Prudente city (BR) for the 61-year period from January 1963 to July 2023. The daily rainfall data were grouped into monthly periods, and the maximum daily rainfall observed for each month was extracted.

The Gumbel distribution demonstrated a good fit to the data for representing maximum rainfall. To illustrate this, we generated a QQ plot, shown in Figure 8 in the Appendix. This plot is constructed by comparing the maximum precipitation observed each month with the estimated values provided by the fitted model. When the dataset closely adheres to the assumed theoretical distribution, the points on the QQ plot will align along a straight line. Thus, the proposed priors for Bayesian analysis and the maximum likelihood approach are discussed for this application.

Estimators with respective variance for both parameters, μ and φ, considering each month of the year are shown in Tables 1 and 2. In addition, the 95% confidence intervals for each parameter are presented in Tables 3 and 4 in Appendix.

Table 1: Point estimation (estimators and standard deviation) of the parameter μ according to the months under priors and MLE.

Months	Jeffreys	Reference	MLE
January	46.7096 (3.2615)	47.5729 (3.7387)	47.2349 (3.2974)
February	45.0349 (3.1456)	44.904 (3.1639)	45.0114 (3.387)
March	31.9277 (2.0745)	31.9566 (2.1516)	32.0212 (2.1439)
April	23.9767 (1.9534)	23.9521 (1.9643)	24.1009 (2.0171)
May	25.3515 (2.1033)	25.601 (2.1123)	25.6014 (2.1468)
June	17.5677 (1.9444)	17.6311 (2.0497)	17.639 (2.0923)
July	9.944 (1.5921)	10.0523 (1.6684)	10.0977 (1.628)
August	9.6276 (1.511)	9.6842 (1.4825)	9.7598 (1.4895)
September	18.6523 (1.6726)	18.5852 (1.7823)	18.6351 (1.7566)
October	33.1487 (1.9727)	33.1306 (2.0811)	33.2378 (2.0478)
November	34.1598 (2.6151)	34.2821 (2.5381)	34.5918 (2.5407)
December	37.7792 (2.3374)	37.5001 (2.3224)	37.871 (2.3928)

Observing Tables 1 and 2, we can verify that the Bayesian estimators obtained with both priors and also under MLE have presented very similar results. This diagnosis corroborates the results obtained from the simulation study due to the moderate sample size (61 years). Furthermore, it is observed that the estimators of the parameters μ and φ are smaller in the driest periods (June, July, August and September) than in the wettest periods. It is possible to understand these results through the expectation of the Gumbel distribution, that is, the expected value of average daily precipitation for drier months will be smaller

Table 2: Point estimation (estimators and standard deviation) of the parameter φ according to the months for the priors and MLE.

Months	Jeffreys	Reference	MLE
January	27.4284 (2.6008)	27.5976 (2.5839)	27.121 (2.4417)
February	28.3781 (2.6937)	28.6375 (2.8157)	28.043 (2.5081)
March	17.8589 (1.7001)	17.9738 (1.7181)	17.633 (1.5875)
April	17.0329 (1.5519)	17.0683 (1.5758)	16.8104 (1.4937)
May	17.9724 (1.6102)	18.1243 (1.6678)	17.7749 (1.5897)
June	17.5586 (1.7558)	17.5882 (1.6821)	17.3233 (1.5494)
July	13.6033 (1.3885)	13.6232 (1.3693)	13.3905 (1.2056)
August	12.2443 (1.2242)	12.3407 (1.2506)	12.0866 (1.103)
September	14.6174 (1.3967)	14.7846 (1.4432)	14.448 (1.3008)
October	16.863 (1.5163)	17.0649 (1.5436)	16.73 (1.5164)
November	20.7403 (2.0058)	20.9183 (1.99)	20.4747 (1.8814)
December	19.5828 (1.8549)	19.8357 (1.8831)	19.4165 (1.7719)

providing smaller μ and φ values. Now, comparing the confidence intervals (MLE) and credibility (Bayesian) presented in Tables 3 and 4 (in the Appendix) for the parameters μ and φ we can see that the intervals are slightly different for both parameters. However, when comparing the credibility intervals obtained under the Jeffreys and Reference priors, there are no differences. As well as the behavior observed for the point estimators, the 95% intervals of the parameters in the driest periods also presented lower limits than in the rainiest periods. In addition, due to the little variability in maximum rainfall, these intervals have presented shorter lengths for the drier months than for the rainy ones.

Return levels are a common measure of extreme events such as in hydrological and climatology applications. Through the inversion of the Gumbel distribution function, we can define the p-th upper quantile as

$$x_p = \mu - \varphi \log\left(-\log\left(1-p\right)\right). \tag{6.1}$$

The value x_p is referred to as the return level. Formally, in the context of extreme rainfall, x_p represents the daily rainfall amount surpassed by the annual maximum each year with a probability of p. Estimates of the return level under the MLE can be easily derived by substituting the estimators for μ and φ into equation (6.1) due to the one-to-one invariance. Under the Bayesian approach, the included an additional step after step 3 in the MCMC procedure where

$$x_{i,p} = \mu_i - \varphi_i \log\left(-\log\left(1-p\right)\right).$$

The chain can be constructed due to the invariance properties of the MCMC and the assumed posterior. There is no need to set additional information or hyperparameters in the MCMC, as it is already constructed for the chains of μ and φ. The credibility intervals are obtained using the percentiles of the obtained chain, and the estimate is the posterior mean.

The estimators obtained by the maximum likelihood and Bayesian methods of the maximum probable daily rainfall for the return times of 2, 4, 10, 20, and 40 years or, respectively, probability levels: 50, 75, 90, 95, and 97.5 percent, determined for the monthly and annual periods are shown in Figure 4. As the Jeffreys and Reference priors have presented similar results from the simulation and Jeffreys prior is invariant under transformations we have used a Bayesian inference with Jeffreys prior to evalue return levels estimators.

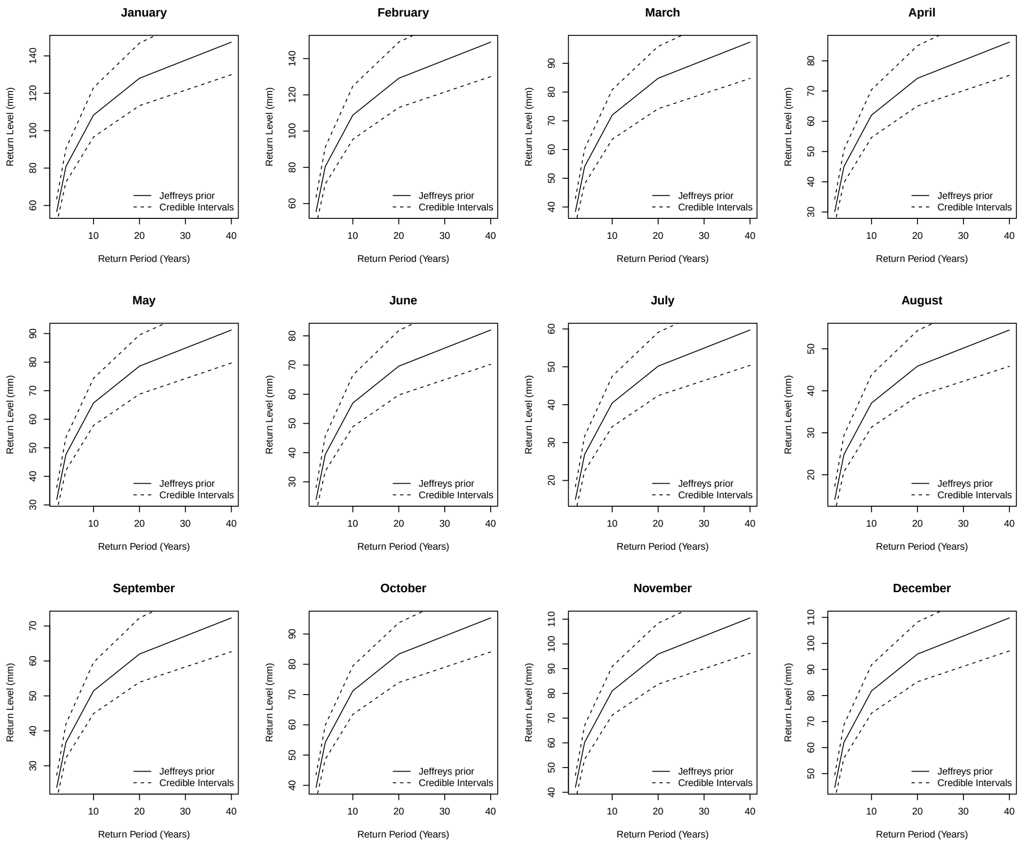

Figure 4: Return level plots for Gumbel distribution of Presidente Prudente annual maximum daily rainfall under Bayesian approach.

Figure 4 illustrates that the Bayesian estimator provides both point estimates and 95% credibility intervals, offering a measure of uncertainty. Notably, these intervals widen as the time interval increases, indicating a growing level of uncertainty associated with the estimations. Moreover, the depicted return levels show a discernible upward trend, with the scale varying from month to month. Particularly during the summer season, the scale tends to be lower than that observed in winter.

7. CONCLUSION

The Gumbel distribution is one of the most commonly used probabilistic tools for calculating extreme events. In this paper, we considered the estimation of the parameters of the Gumbel distribution using the Bayesian approach under objective priors. Objective priors can be determined immediately to carry out a Bayesian analysis for the model of interest, as there are several such proposed priors in the literature. Furthermore, informative priors, based on subjective methods, are often more challenging to apply in practice. Therefore, in this paper, the main objective priors have been derived for the Gumbel distribution such as Jeffreys, MDI, Tibshirani, and reference priors. Some of these priors have not been considered in Bayesian analysis under Gumbel distribution applications in the literature, due to their similar results from the simulation study.

This study has provided a general form for these objective priors, establishing a method capable of being updated automatically without the need for expert intervention. It allows for quick and straightforward performance of Bayesian analysis from both the theoretical and computational points of view. Also, we have observed that the reference, MDI, and Tibshirani priors have the same form. Therefore, the obtained posterior has excellent properties, such as one-to-one invariance, consistent marginalization, consistent sampling properties, and coverage probabilities, making it easier to obtain the posterior of these functions through the MCMC algorithm. Although Jeffreys' prior differs from these three priors, it performs quite similarly to Jeffreys regardless of the sample size, based on the simulation study. On the other hand, the reference prior provided more accurate credibility intervals.

Importantly, it has also been shown that all studied improper priors in this paper lead to proper posterior distributions, which does not occur in many other distributions such as the Generalized Exponential, Gamma, Generalized Gamma, Generalized Extreme Value (GEV), and other distributions (see for instance, Ramos et al., 2023). Northrop and Attalides (2016) and Zhang and Shaby (2023) considered such priors for a generalized version, but the priors obtained differ when reduced to the Gumbel distribution. Moreover, the posterior obtained with the Jeffreys priors and the MDPI returns improper posteriors, which prevents the posterior distribution from being a probability distribution, and therefore it cannot be used in Bayesian inference. In contrast, our approach yields a proper posterior with prior distributions of the form $\pi(\mu, \varphi) \propto \varphi^{-k}$ for $k \geq 1$ for sample sizes $n \geq 2$. The reference priors for the GEV model depend on the order of the parameters, and more importantly, they do not align with ours when the generalized model is reduced to the Gumbel distribution. Hence, our results cannot be viewed as particular cases within more general models.

The estimators of the parameters of the Gumbel distribution obtained by the maximum likelihood method are slightly different from the Bayesian estimators, but those obtained by the Bayesian analysis provided more accurate estimates, primarily for small sample sizes. Overall, the Gumbel distribution was successfully applied to analyze the data of maximum monthly rainfall, effectively reproducing the rainfall regime in the Presidente Prudente region. There are a large number of possible extensions to this current work. The presence of censoring, covariates, and long-term survival are common in practice. Our approach should be investigated further in these contexts.

ACKNOWLEDGMENTS

Nixon Jerez-Lillo was funded by the National Agency for Research and Development (ANID)/Scholarship Program/Doctorado Nacional/2021-21210981.

CODE AVAILABILITY

All the functions and procedures concerning implementation included in the article have been implemented in R Core Team. The codes will be available as `https://github.com/njerezlillo/objectivepriorsgumbel/tree/main`.

REFERENCES

Berger, J. O., Bernardo, J. M., et al. (1992). On the development of reference priors. *Bayesian statistics*, 4(4):35–60.

Berger, J. O., Bernardo, J. M., Sun, D., et al. (2015). Overall objective priors. *Bayesian Analysis*, 10(1):189–221.

Bernardo, J. M. (1979). Reference posterior distributions for bayesian inference. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 113–147.

Bernardo, J. M. (2005). Reference analysis. *Handbook of statistics*, 25:17–90.

Bickel, P. J. and Doksum, K. A. (2015). *Mathematical statistics: basic ideas and selected topics, volumes I-II package*. CRC Press, New York.

Chbab, E., Van Noortwijk, J., and Duits, M. (2000). Bayesian frequency analysis of extreme river discharges. In *Proceedings of International Symposium on Flood Defence*, Kassel.

Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335.

Coles, S. (2003). The use and misuse of extreme value models in practice. In *Extreme values in finance, telecommunications, and the environment*, pages 98–119. Chapman and Hall/CRC, Boca Raton.

Consonni, G., Fouskakis, D., Liseo, B., Ntzoufras, I., et al. (2018). Prior distributions for objective bayesian analysis. *Bayesian Analysis*, 13(2):627–679.

Cox, D. R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–39.

Elkahlout, G. (2006). Bayes estimators for the extreme-value reliability function. *Computers & Mathematics with Applications*, 51(3-4):673–679.

Ferreira, P. H., Ramos, E., Ramos, P. L., Gonzales, J. F., Tomazella, V. L., Ehlers, R. S., Silva, E. B., and Louzada, F. (2020). Objective bayesian analysis for the lomax distribution. *Statistics & Probability Letters*, 159:108677.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. *Bayesian statistics*, 4:641–649.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 186, pages 453–461. The Royal Society.

Kass, R. E. and Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91(435):1343–1370.

Kotz, S. and Nadarajah, S. (2000). *Extreme value distributions: theory and applications*. world scientific, London.

Lye, L. M., Hapuarachchi, K., and Ryan, S. (1993). Bayes estimation of the extreme-value reliability function. *IEEE transactions on Reliability*, 42(4):641–644.

Mousa, M. A., Jaheen, Z., and Ahmad, A. (2002). Bayesian estimation, prediction and characterization for the gumbel model based on records. *Statistics: A Journal of Theoretical and Applied Statistics*, 36(1):65–74.

Nadarajah, S. and Choi, D. (2007). Maximum daily rainfall in south korea. *Journal of Earth System Science*, 116:311–320.

Neal, P. and Roberts, G. (2008). Optimal scaling for random walk metropolis on spherically constrained target densities. *Methodology and Computing in Applied Probability*, 10:277–297.

Northrop, P. and Attalides, N. (2016). Posterior propriety in bayesian extreme value analyses using reference priors. *Statistica Sinica*, 26(2).

Onwuegbuche, F. C., Affognon, S. B., Enoc, E. P., Akinade, M. O., et al. (2019). Application of extreme value theory in predicting climate change induced extreme rainfall in kenya. *International Journal of Statistics and Probability*, 8(4):85–94.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ramos, P. L., Almeida, M. H., Louzada, F., Flores, E., and Moala, F. A. (2022). Objective bayesian inference for the capability index of the weibull distribution and its generalization. *Computers & Industrial Engineering*, 167:108012.

Ramos, P. L., Mota, A. L., Ferreira, P. H., Ramos, E., Tomazella, V. L., and Louzada, F. (2021). Bayesian analysis of the inverse generalized gamma distribution using objective priors. *Journal of Statistical Computation and Simulation*, 91(4):786–816.

Ramos, P. L., Rodrigues, F. A., Ramos, E., Dey, D. K., and Louzada Neto, F. (2023). Power laws distributions in objective priors. *Statistica Sinica*, 33(3):1–51.

Rostami, M. and Adam, M. B. (2013). Analyses of prior selections for gumbel distribution. *MATEMATIKA: Malaysian Journal of Industrial and Applied Mathematics*, pages 95–107.

Tiao, G. C. and Box, G. E. (1973). Some comments on "bayes" estimators. *The American Statistician*, 27(1):12–14.

Tibshirani, R. (1989). Noninformative priors for one parameter of many. *Biometrika*, 76(3):604–608.

Tomazella, V. L. D., Jesus, S. R. d., Louzada Neto, F., Nadarajah, S., and Ramos, P. L. (2020). Reference bayesian analysis for the generalized lognormal distribution with application to survival data. *Statistics and Its Interface*, 13(1):139–149.

Vidal, I. (2014). A bayesian analysis of the gumbel distribution: an application to extreme rainfall data. *Stochastic environmental research and risk assessment*, 28:571–582.

Yılmaz, A., Kara, M., and Özdemir, O. (2021). Comparison of different estimation methods for extreme value distribution. *Journal of Applied Statistics*, 48(13-15):2259–2284.

Zellner, A. (1977). Maximal data information prior distributions. *In New Methods in the applications of Bayesian Methods*.

Zellner, A. (1984). Maximal data information prior distributions. *Basic Issues in Econometrics*.

Zhang, L. and Shaby, B. A. (2023). Reference priors for the generalized extreme value distribution. *Statistica Sinica*, 33:2185–2208.

A. BOXPLOTS

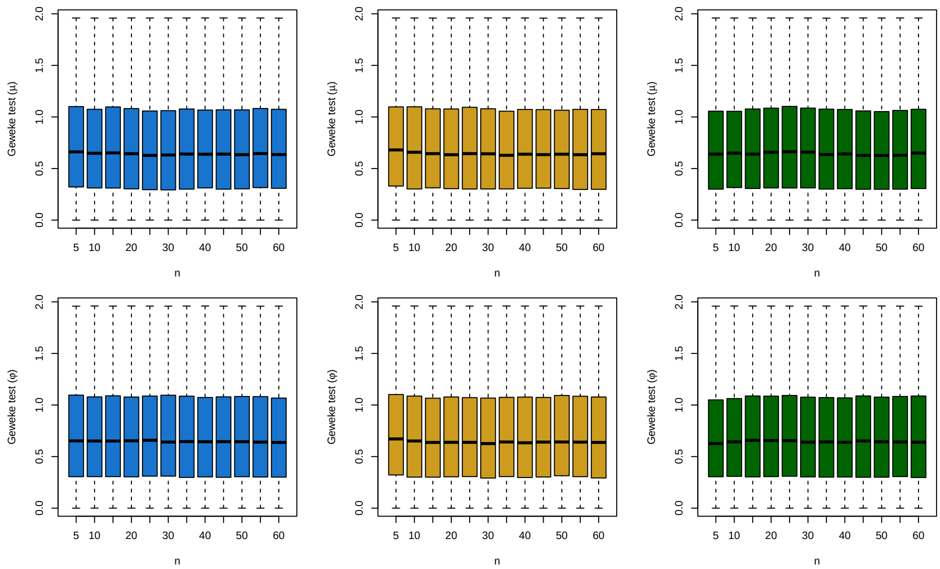

Figure 5: Boxplots of Geweke test results for chains generated for μ (first row) and φ (second row) in Scenario 1, across different sample sizes. The columns correspond to the prior types: Jeffreys (left), reference (middle), and informative (right).

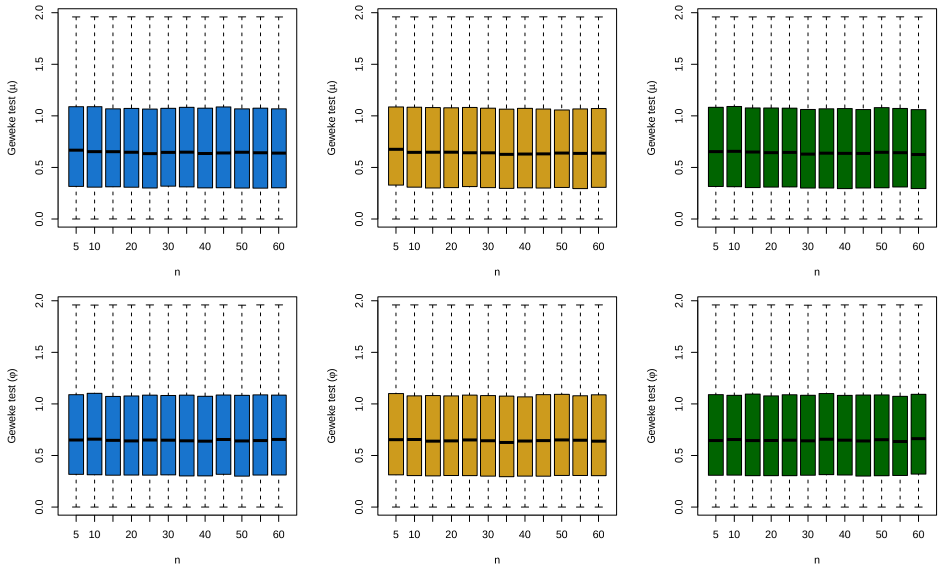

Figure 6: Boxplots of Geweke test results for chains generated for μ (first row) and φ (second row) in Scenario 2, across different sample sizes. The columns correspond to the prior types: Jeffreys (left), reference (middle), and informative (right).

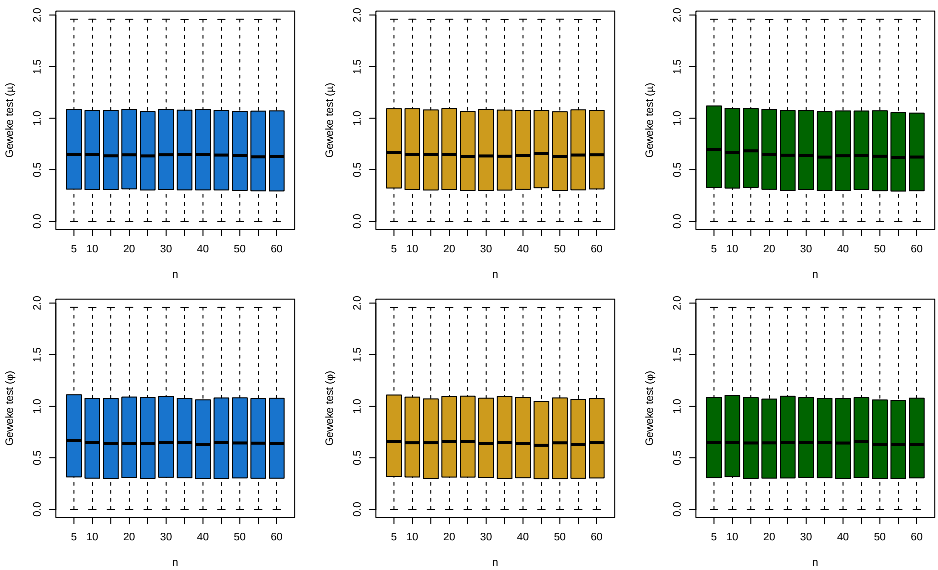

Figure 7: Boxplots of Geweke test results for chains generated for μ (first row) and φ (second row) in Scenario 3, across different sample sizes. The columns correspond to the prior types: Jeffreys (left), reference (middle), and informative (right).

B. QQ PLOTS

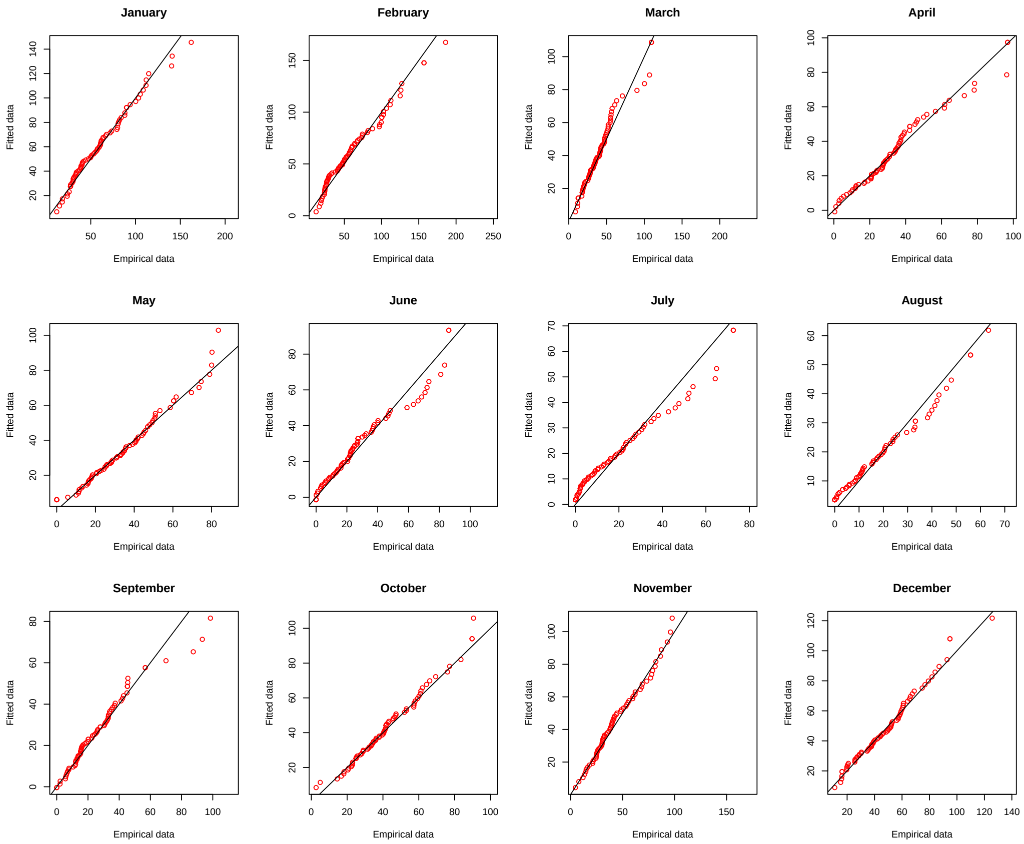

Figure 8: QQ plots comparing the observed maximum precipitation each month to the estimated values provided by the fitted Gumbel model.

C. 95% CONFIDENCE INTERVALS

Table 3: 95% Interval estimation of the parameter μ, according to the months under priors and MLE.

Months	Jeffreys	Reference	MLE
January	(40.7292; 52.8546)	(40.4019; 53.1493)	(40.7719; 9.7604)
February	(37.9284; 52.7275)	(37.9918; 53.6106)	(38.3728; 10.0256)
March	(27.5315; 36.2134)	(27.6261; 36.0207)	(27.8192; 6.3458)
April	(19.9457; 27.9605)	(20.0557; 27.8602)	(20.1474; 5.9707)
May	(21.1885; 29.3611)	(21.2016; 29.336)	(21.3936; 6.3547)
June	(13.546; 21.3465)	(13.5243; 21.5384)	(13.5381; 6.1932)
July	(6.7146; 13.0758)	(6.8429; 13.0597)	(6.9067; 4.819)
August	(6.7487; 12.5928)	(6.7724; 12.6878)	(6.8403; 4.409)
September	(15.1897; 21.6401)	(15.1324; 21.8419)	(15.1921; 5.1996)
October	(29.0812; 36.974)	(29.1051; 37.1372)	(29.2242; 6.0614)
November	(29.3227; 38.9873)	(29.4344; 39.0837)	(29.612; 7.5204)
December	(33.0443; 42.1171)	(33.1589; 42.2736)	(33.1811; 7.0827)

Table 4: 95% Interval estimation of the parameter φ, according to the months under priors and MLE.

Months	Jeffreys	Reference	MLE
January	(22.9506; 33.3489)	(23.1399; 33.8644)	(22.3352; 7.2276)
February	(23.4822; 34.1252)	(23.5924; 34.5519)	(23.1271; 7.4239)
March	(14.805; 21.3189)	(14.7842; 21.6177)	(14.5215; 4.6991)
April	(14.2517; 20.2367)	(14.3139; 20.2949)	(13.8828; 4.4213)
May	(15.1106; 21.0789)	(15.1129; 21.4296)	(14.659; 4.7056)
June	(14.5891; 21.336)	(14.6928; 21.4387)	(14.2866; 4.5861)
July	(11.1347; 16.3576)	(11.2307; 16.5023)	(11.0276; 3.5685)
August	(10.0519; 14.7935)	(10.2132; 14.9868)	(9.9248; 3.2648)
September	(12.1823; 17.6832)	(12.2646; 17.9292)	(11.8985; 3.8503)
October	(14.1854; 20.1829)	(14.2351; 20.2954)	(13.7579; 4.4885)
November	(17.1318; 24.816)	(17.384; 25.3427)	(16.7872; 5.5689)
December	(16.3583; 23.5061)	(16.4848; 23.7801)	(15.9436; 5.2448)