mathematics

MDPI

Article

Bayesian Variable Selection in Generalized Extreme Value Regression: Modeling Annual Maximum Temperature

Jorge Castillo-Mateo [1,*], Jesús Asín [1], Ana C. Cebrián [1], Jesús Mateo-Lázaro [2] and Jesús Abaurrea [1]

[1] Department of Statistical Methods, University of Zaragoza, 50009 Zaragoza, Spain
[2] Department of Earth Sciences, University of Zaragoza, 50009 Zaragoza, Spain
* Correspondence: jorgecm@unizar.es

Abstract: In many applications, interest focuses on assessing relationships between covariates and the extremes of the distribution of a continuous response. For example, in climate studies, a usual approach to assess climate change has been based on the analysis of annual maximum data. Using the generalized extreme value (GEV) distribution, we can model trends in the annual maximum temperature using the high number of available atmospheric covariates. However, there is typically uncertainty in which of the many candidate covariates should be included. Bayesian methods for variable selection are very useful to identify important covariates. However, such methods are currently very limited for moderately high dimensional variable selection in GEV regression. We propose a Bayesian method for variable selection based on a stochastic search variable selection (SSVS) algorithm proposed for posterior computation. The method is applied to the selection of atmospheric covariates in annual maximum temperature series in three Spanish stations.

Keywords: climate change; extreme value analysis; Markov chain Monte Carlo; non-stationary; stochastic search variable selection

MSC: 62F15; 62G32; 62H12; 62P12

Citation: Castillo-Mateo, J.; Asín, J.; Cebrián, A.C.; Mateo-Lázaro, J.; Abaurrea, J. Bayesian Variable Selection in Generalized Extreme Value Regression: Modeling Annual Maximum Temperature. *Mathematics* **2023**, *11*, 759. https://doi.org/10.3390/math11030759

Academic Editor: Manuel Alberto M. Ferreira

Received: 24 December 2022
Revised: 28 January 2023
Accepted: 29 January 2023
Published: 2 February 2023

1. Introduction

Many studies have shown that climate change not only yields an increase in mean temperature but also in extreme weather conditions; see Rahmstorf and Coumou [1], the IPCC report, [2] and references therein. However, the effect of climate change in temperature is not the same in the mean as in the extreme quantiles of the distribution [3,4]. Extreme temperatures, although rare by definition, can have large impacts on economy, ecosystems, and human health [5,6]. These impacts are larger than those provoked by changes in the mean [7,8].

The aim of this work is to develop a new Bayesian variable selection method in the generalized extreme value (GEV) modeling framework to study extreme events. This methodology must be able to identify automatically promising subsets of important covariates from a large set of potential covariates. The main ideas of the methodology are based on embedding a Bayesian stochastic search variable selection (SSVS) method within the GEV model specification for the data. We implement an adaptive Metropolis-within-Gibbs Markov chain Monte Carlo (MCMC) algorithm to obtain posterior samples from the model parameters. The usefulness of the proposed method is shown in an application to the analysis of extremes in three Spanish series of daily maximum temperature as a function of atmospheric covariates.

Statistical analysis of extremes is intrinsically difficult, since they occur with a lower frequency than events around the central part of the distribution. The branch of statistics called extreme value analysis (EVA) aims to propose specific asymptotic distributions for this type of observations. The most common approaches in EVA are the threshold

exceedance models, where extreme events are defined as those that exceed some high threshold and block maxima analysis. The latter, that will be used in this work, considers the maximum of the process over n time units of observation, and then models the block maxima. A fundamental result of this theory for independent and identically distributed (i.i.d.) random variables is that if there is a non-trivial limiting distribution for normalized maxima of n of these random variables, then the limit as n goes to infinity is a GEV distribution.

It is common to assume that annual maximum temperatures follow a stationary GEV distribution [9,10]. However, to study the effects of global warming, one should consider models that allow capturing the evolution of maximum temperatures over time, such as non-stationary GEV models. A common technique to consider a non-stationary behavior in a parametric model is to regress the parameters on known covariates. In this way, and using temporal trends or influential time-varying covariates, one can model how annual maximum temperatures are changing over time. In this framework, Gao and Zheng [11] analyze the relationship between annual maximum temperature in Chinese stations and climate indices of wide spatial effect as Arctic Oscillation. The advantages of non-stationary models versus stationary models are shown in several works [12,13] and in the comparative analysis of models for annual maximum temperatures [14].

Many approaches have been proposed for parameter estimation in GEV models; e.g., techniques based on versions of probability plots, moment-based estimators, estimation based on order statistics, and likelihood-based methods [15]. A potential issue with using likelihood methods for the GEV concerns the regularity conditions that are needed to guarantee the usual asymptotic properties of maximum likelihood estimators. Those conditions do not hold in the GEV model because the end-points of the GEV distribution are functions of the parameters and, consequently, the standard asymptotic likelihood properties are not automatically applicable [16]. Other estimation methods are penalized maximum likelihood [17], as well as several Bayesian approaches [8,18,19]. Friederichs and Thorarinsdottir [20] compare different GEV estimation methods such as maximum likelihood estimation, optimum score estimation with the continuous ranked probability score, and Bayesian estimation. They found that the Bayesian approach yields the highest overall prediction skills.

One of the most important research lines in climate change studies is to obtain projections under different climate change scenarios of the atmospheric covariates with a strong impact on human activity, such as precipitation [21] or temperature extremes [22]. Earth system models (ESMs) provide projections of temperature under different scenarios, but they have two main limitations. First, they work at a large spatial scale, so that physical or statistical downscaling methods are needed to obtain temperature projections at local scale [23]. Second, daily temperature projections for current climate scenarios show a good reproduction of temperature in the central part of its distribution, but it is not adequate in extreme values. Therefore, specific statistical downscaling procedures for daily temperature extremes at a local scale have to be used, both when extremes are defined as excesses over threshold [24,25] or as maximum values [26,27].

Statistical downscaling methods for extreme temperature should include atmospheric time-varying covariates that capture the non-stationary behavior of extremes due to the effect of climate change. The number of potential covariates is high, because ESMs provide information, in surface and different pressure levels, about temperature, humidity, geopotential, and other variables. In addition, a high collinearity is found in the potential large scale predictors [28]. Our set of potential covariates is formed by all the atmospheric variables available from CMIP6 climate models, with the only restriction being that they have to adequately represent the observed climate [29,30].

In this context, interest focuses on selecting the atmospheric covariates among the high dimensional set of potential variables that could be included in a statistical downscaling model for annual maximum temperatures. A wide variety of selection procedures based on the comparison of all 2^k possible submodels has been proposed in terms of the value

of a goodness-of-fit or error measure, such as AIC, BIC, or Mallows's C_p. However, if k is large, the computational burden of these procedures is too high, especially if estimation is carried out in a Bayesian framework. To mitigate the computational costs, heuristic methods such as stepwise procedures are usually applied to select a smaller number of potential variables. The covariate selection problem appears in many regression-type models [31], but most of them focus on normal regression models and maximum likelihood estimation. Bayesian methods for variable selection can be very useful to find a subset of potential explanatory variables to be included in a final model, given a large set of k potential covariates. However, some caution must be taken because the usual model selection by Bayes factor is known to be sensitive to hyperparameter choice in hierarchical models [32]. In this general framework, George and McCulloch [33] proposed the Bayesian SSVS method to reduce the number of promising models for further consideration. The idea is to embed the regression setup in a Bayesian hierarchical normal mixture model that avoids the calculation of the posterior probabilities of all 2^k subsets. Efforts are made to develop software implementing selection tools, e.g., the R packages **spikeSlabGAM** [34] and **gamselBayes** [35] use SSVS, but they do not incorporate GEV regression. Yu et al. [36] develop Bayesian methods for variable selection with a simple and efficient SSVS algorithm in quantile regression. There are some covariate selection methods for GEV parametric models using a Bayesian estimation approach. Friederichs and Thorarinsdottir [20] perform a variable selection procedure using a method based on a reversible jump MCMC algorithm. El Adlouni and Ouarda [37] develop a birth–death MCMC procedure to carry out both parameter estimation and Bayesian model selection.

As an alternative to the above, we propose a Bayesian SSVS method for variable selection in GEV regression models. This method is particularly useful in the development of downscaling methods for maximum temperatures. Thus, we provide a modeling approach that allows us not only to identify atmospheric variables that have an influence on annual maximum temperatures, but also a statistical downscaling approach to generate projections of the annual maximum temperature under different climate change scenarios. The Bayesian hierarchical model can be used to simulate series of annual maximum temperature. This simulation tool could be used, e.g., to study and make inference on different types of features of these series such as the occurrence and magnitude of its record events [38,39], in the observed period and future climate scenarios, using the ESM projections of the covariates.

Finally, note that it is essential to handle uncertainty based on the available evidence [40] and accounting for the substantial uncertainty that occurs in the variable selection process is needed [36]. Bayesian approaches provide a convenient paradigm for accommodating uncertainty in model selection [41]. Prior distributions are expressions of the prior knowledge and uncertainty about the distribution of the parameters. Bayes' theorem combines these priors with the model and data to obtain the posterior distribution, which summarizes updated knowledge and uncertainty about the parameters. Briefly, SSVS specifies a prior distribution with binary inclusion indicators in the covariates. Subsequently, using MCMC algorithms, it samples thousands of regression models in order to characterize the model uncertainty regarding both the covariate set and the regression parameters by implementing local changes to a single covariate at a time. In this way, SSVS is a standard way of handling model uncertainty problems [42].

The remainder of the paper is organized as follows. Section 2 describes temperature data and atmospheric covariates. It also describes the proposed GEV regression SSVS (GEV-SSVS) algorithm. Section 3 presents a simulation study and the results of applying the GEV-SSVS approach to three Spanish temperature series. Main conclusions and future work are summarized in Section 4. The method proposed in this paper is available in the R package **GEVSSVS** through GitHub at https://github.com/JorgeCastilloMateo/GEVSSVS (accessed on 10 December 2022).

2. Materials and Methods

2.1. Materials

Three Spanish series of daily maximum temperature obtained from the European Climate Assessment & Dataset (ECA&D) [43] are analyzed. The series correspond to Barcelona (Fabra Observatory), Madrid (Retiro), and Zaragoza (Airport); their location within the Iberian Peninsula is shown in Figure 1. They have different orography and represent different climates. Barcelona is located in the northeastern part of the Mediterranean coast at the Iberian Peninsula, with a maritime Mediterranean climate. July and August are the warmest months, with average temperatures around 28 and 22 °C during the day and night, respectively. Madrid lies in the center of the Iberian Peninsula on the southern Central Plateau, at 657 m a.s.l. It has a Mediterranean climate with continental and semi-arid influences. Summers are hot, dry, and sunny. July is the warmest month, with average temperatures during the day around 33 °C, and daily maxima commonly climbing over 35 °C. Zaragoza lies at 243 m a.s.l. in the northeast of the Iberian Peninsula, in the rather arid depression formed by the Ebro Valley. It has a Mediterranean semi-arid climate. The summer lasts from early June to early September, with an average daily maximum temperature over 28 °C. The hottest month of the year in Zaragoza is July.

Figure 1. Location of the temperature series in the Iberian Peninsula.

The three series are recorded up to 2021, but they have different lengths: Barcelona starts in 1914, Madrid at 1920, and Zaragoza at 1951. No missing or suspect values are observed in Barcelona, 145 missing and 5 suspect values are observed in Madrid, and 3 missing values in Zaragoza. However, none of these values occur in the summer period, where the annual maximum is observed; consequently, the three series of annual maximum temperatures do not have any missing value.

The set of potential covariates to be included in the model are selected from the atmospheric variables from reanalysis ERA-20C from the European Centre for Medium-Range Weather Forecasts (ECMWF) [44]. In this work, we consider temperature (°K), specific humidity (kg/kg,) and geopotential (m^2s^{-2}) at 850, 700, 500, and 300 hPa pressure levels. The monthly mean of these variables is available in a grid $1^\circ \times 1^\circ$ covering the Iberian Peninsula. Then, we consider 48 potential covariates, which correspond to the monthly mean in July (the hottest month in the three study series) of the 12 variables in the four points of the spatial grid surrounding the station under study. The 12 variables above are denoted by t, q and z, respectively; a suffix specifying the level; e.g., $q.500$ denotes the specific humidity at 500 hPa. Since the series of the covariates are available from 1900 to

2010, all the analyses below consider the series from the starting year of the temperature series up to 2010.

We will also consider the temperature (°K) at 2 m over ground level, denoted by *t2m*. This covariate explains much of the variability of the series, but it is not well reproduced by the ESM in the current climate scenario; consequently, it is not suitable to be used in models to generate projections. By this reason, it is used as an illustration in a preliminary model but not in the final model. The R package **ecmwfr** [45] was used to download the data from the ECMWF (accessed on 1 April 2022).

2.2. Methods

This section proposes a modeling approach for non-stationary annual maximum temperatures based on the GEV distribution, and presents the Bayesian variable selection method. The basic structure of GEV-SSVS modeling and algorithm are illustrated in Figure 2 and its components are described in detail below. On the left, the process selection workflow shows the steps from the database and pre-selection to the final selected model. The right part of the figure outlines the MCMC algorithm that is developed to fit the GEV-SSVS models.

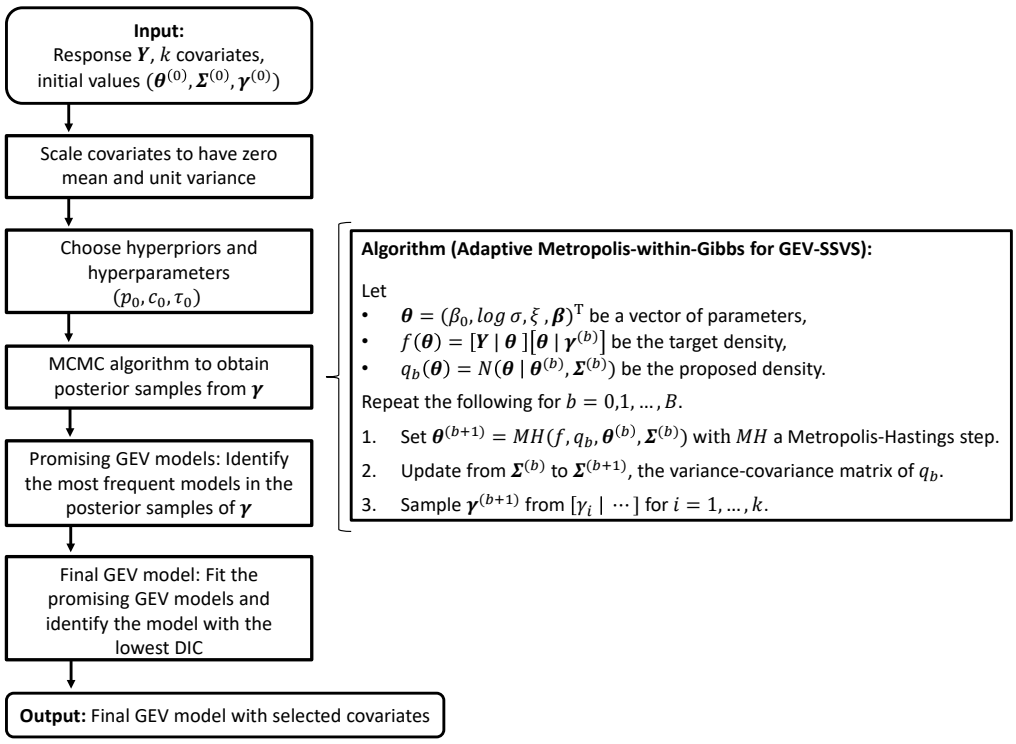

Figure 2. The framework of the GEV-SSVS method. **Left**: outline of the workflow. **Right**: outline of the iterative MCMC algorithm to obtain posterior samples of the model parameters.

2.2.1. Modeling Annual Maximum Temperature

The aim of this work is to model annual maximum temperature: to characterize the statistical behavior of the random variables that represent the maximum of a process over n time units of observation,

$$M_n = \max\{Z_1, \ldots, Z_n\},$$

where $Z_1, \ldots, Z_n$ is a sequence of random variables. In our case since n is the number of observations in a year, M_n corresponds to the annual maximum.

The distribution of M_n is difficult to obtain in a general case. However, EVA provides useful asymptotic results to characterize the GEV distribution as its limit. We remind that the $GEV(\mu, \sigma, \xi)$ cumulative distribution function is defined as

$$G(z) = \begin{cases} \exp\left\{-\left[1+\xi\frac{z-\mu}{\sigma}\right]_+^{-1/\xi}\right\}, & \text{if } \xi \neq 0, \\ \exp\left\{-\exp\left[-\frac{z-\mu}{\sigma}\right]\right\}, & \text{if } \xi = 0, \end{cases} \tag{1}$$

where $x_+ = \max\{0, x\}$, the scale parameter is $\sigma > 0$, and the location and shape parameters are $-\infty < \mu, \xi < \infty$, respectively. This distribution contains three distributions as particular cases: the heavy-tailed Fréchet distribution is obtained with $\xi > 0$, the upper bounded Weibull distribution with $\xi < 0$, and the Gumbel distribution with $\xi = 0$.

Under quite general conditions, for $Z_1, \ldots, Z_n$ i.i.d. random variables, if there exist two sequences $a_n > 0$ and b_n such that the following limit converges to a non-degenerate distribution function $G(z)$,

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{M_n - b_n}{a_n} \leq z\right) = G(z),$$

then $G(z)$ is a GEV distribution [46]. With this result, the GEV distribution is the only possible limit of a linear normalization of the maximum of a sequence of i.i.d. random variables. This justifies the use of the GEV distribution for modeling the distribution of maxima of long sequences. Clearly, when the distribution of the normalized M_n can be approximated by a GEV distribution, the maximum M_n itself can also be approximated by a different member of the same family and, in any case, the parameters have to be estimated from the data.

Note that the previous result assumes an underlying process of i.i.d. random variables. However, in a certain sense and subject to specified limitations, the GEV model is still applicable under short-term temporal dependence, as is often the case with meteorological data [47].

Modeling Non-Stationarity

In most environmental processes, and temperature in particular, stationarity is not a reasonable assumption due to the existence of seasonal effects and trends linked to climate change or other effects. In a non-stationary time process, the behavior of the variables changes systematically over time, and it cannot be modeled by a sequence of i.i.d. variables. This departure from the assumption of the GEV characterization as a limit distribution of maxima does not allow to guarantee the validity of the approximation. There is no general theory for characterizing the distribution of maxima in non-stationary processes. Results are available for very specific forms of non-stationarity, but often they are not useful for describing the non-stationary behavior of real processes. However, it is usually the use of the GEV model as a basic template that is enhanced by statistical modeling, as suggested by Coles [15].

To sum up, in a non-stationary framework, there are asymptotic arguments that support the use of the GEV distribution for modeling the maximum in any year, but the existence of a trend or the influence of a time-varying covariate in the process makes a model which assumes a constant distribution through time unsuitable. Then, a reasonable approach to model $\mathbf{Y}$, the series of variables Y_t that represent the annual maximum temperature in year t, is to assume that they are independent variables with distribution

$$Y_t \mid \mu_t, \sigma, \xi \sim GEV(\mu_t, \sigma, \xi), \qquad t = 1, \ldots, T, \tag{2}$$

with $\mu_t = h(\mathbf{x}_t, \beta_0, \boldsymbol{\beta})$, where h is a specified function, $\mathbf{x}_t$ is a vector of covariates at time t, β_0 is an intercept, $\boldsymbol{\beta}$ is a vector of parameters, and T is the length of the series. Similar time-varying structures can be considered also in σ and ξ, the scale and shape parameters.

In most real applications, it is assumed that the variations through time in the process can be captured just by modeling the location parameter μ as a time-varying function. However, this assumption should be checked.

The next step in the modeling process is to specify the function h. Given a set of time-varying covariates $\mathbf{x}_t^\top = (x_{t1}, \ldots, x_{tk})$, we consider the following function

$$\mu_t = \beta_0 + \mathbf{x}_t^\top \boldsymbol{\beta}, \qquad t = 1, \ldots, T, \tag{3}$$

where β_0 is the intercept, and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)^\top$ are the regression coefficients. Note that the vector of time-varying covariates may include variables such as a linear or polynomial trend across years. The stationary model, where the distribution of annual temperature maxima is constant over time, is a special case of this non-stationary model with $\boldsymbol{\beta} = \mathbf{0}$.

2.2.2. Bayesian Variable Selection, Model Fitting, and Prediction

The proposed model is estimated in a Bayesian framework. Although Bayesian estimation of GEV parameters has been studied for a long time, selecting an appropriate model becomes an important issue when there is a large number of potential covariates. Since the basic principle is parsimony (that is, to obtain the simplest model that fits adequately the data) we have to assess the strength of evidence of a more complex model structure. However, if k covariates are available, computation of the posterior probabilities of the 2^k possible subsets of covariates is computationally expensive and methods that provide some preliminary selection will be clearly helpful.

Our aim is to provide a method to select subsets of promising covariates for further study in a parametric GEV model. We propose to combine the GEV model with the SSVS algorithm, adapted to the case of selecting the regression coefficients of the linear predictor that defines one parameter of a GEV distribution. The underlying idea of the SSVS method is to identify as sets of promising covariates those with higher posterior probability. To that end, it uses MCMC to indirectly sample from the multinomial posterior distribution on the set of possible subsets and identify the higher probability sets by their more frequent appearance in the MCMC iterations.

Bayesian Variable Selection

To implement the SSVS method, the entire regression setup is embedded in a hierarchical Bayes normal mixture model, where latent variables are used to identify subset choices. More precisely, we introduce the latent variables $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_k)^\top$ to represent the normal mixture by

$$\beta_i \mid \gamma_i \sim (1 - \gamma_i) N(0, \tau_i^2) + \gamma_i N(0, c_i^2 \tau_i^2), \quad i = 1, \ldots, k,$$

where each γ_i is a binary variable that can take values 0 or 1 with $\mathrm{P}(\gamma_i = 1) = p_i$. In this way,

$$\beta_i \sim \begin{cases} N(0, \tau_i^2) & \text{if } \gamma_i = 0, \\ N(0, c_i^2 \tau_i^2) & \text{if } \gamma_i = 1. \end{cases}$$

This means that setting a small $\tau_i > 0$, if $\gamma_i = 0$, then β_i could be safely estimated by zero. In addition, setting a large $c_i > 1$, if $\gamma_i = 1$, a non-zero estimate of β_i should be included in the model. Specific choices of τ_i and c_i for this purpose are given below. In this setup, p_i may be thought of as the prior probability that β_i will require a non-zero estimate, and consequently that the ith covariate should be included in the model. Summarily, the posterior distribution of the vector $[\boldsymbol{\gamma} \mid \boldsymbol{Y}]$ will give higher probabilities to the model with the most promising covariates.

To complete the model, the values of c_i and τ_i are specified, following [48]. The approach is based on considering $\delta_i > 0$, the intersection point of the densities of $N(0, \tau_i^2)$ and $N(0, c_i^2 \tau_i^2)$. Note that $|\beta_i| \leq \delta_i$ corresponds to the region where $N(0, \tau_i^2) \succcurlyeq N(0, c_i^2 \tau_i^2)$, and $|\beta_i| > \delta_i$ corresponds to the region where $N(0, \tau_i^2) \preccurlyeq N(0, c_i^2 \tau_i^2)$, where $\succcurlyeq$ and $\preccurlyeq$

denote density dominance. Large posterior probability of γ under $[\gamma \mid \boldsymbol{Y}]$ suggests that $|\beta_i| \leq \delta_i$ for $\gamma_i = 0$, and $|\beta_i| > \delta_i$ for $\gamma_i = 1$. Thus, c_i and τ_i should be selected so that if $|\beta_i| \leq \delta_i$ it would be preferable to set $\beta_i = 0$ and exclude the ith covariate from the model. Once δ_i has been selected, choosing c_i between 10 and 100 is recommended to achieve an effective separation between both densities. For a given c_i and δ_i, the implicit value of τ_i is then obtained as

$$\tau_i = \left[2\log(c_i)c_i^2/(c_i^2-1)\right]^{-1/2}\delta_i.$$

Note that δ_i could be defined considering that it expresses a cut-off value to define below it an effect close to zero and therefore negligible. This value can be based on a previous exploratory analysis or comparing results on a grid of values.

Prior Specification and Model Fitting

For homogeneity and simplicity we consider common values of $p_0 \equiv p_i$, $c_0 \equiv c_i$, $\tau_0 \equiv \tau_i$, and $\delta_0 \equiv \delta_i$ $(i = 1, \ldots, k)$ across covariates. The number of covariates included in the final model will be sensitive to the value of p_0. Consequently, p_0 can be fixed considering previous knowledge or analyses about the number of covariates in a final model. For our data we expect the final model to use no more than four or five covariates, so we set the probabilities of covariate inclusion to $p_0 = 1/10$. A more general approach could estimate p_0 from the data, and in this case a mathematically convenient prior would be $p_0 \sim Beta(a_0, b_0)$.

It is appropriate to assign the intercept a diffuse prior $\beta_0 \sim N(0, 10^2)$, and analogously for the scale parameter $\log\sigma \sim N(0, 2)$. As suggested in the literature, we specify a weakly informative prior for the shape parameter $\xi \sim N(0, 1)$. Here, weakly informative means that the prior has a middling amount of certainty, being neither too diffuse nor too restrictive, because it is generally difficult to estimate ξ from the data. The posterior inference is not very sensitive to these prior distributions.

Before fitting the model, it is convenient to scale the covariates by mean and variance. We implement a sensitivity analysis to understand the effect of c_0 and δ_0 in the final model. Based on preliminary analyses fitting different models with different parameters, we set $c_0 = 100$ and an intersection point $\delta_0 = 0.15$, or equivalently $\tau_0 = 0.05$. Note that δ_0 defines the intersection point between the normal distributions that express the significant or insignificant effect of the scaled covariate.

Using this hierarchical model, MCMC methods are used to obtain samples from the joint posterior distribution. The sampling algorithm is an adaptive Metropolis-within-Gibbs version. The binary variables are sampled from their full conditional distribution one-by-one from

$$[\gamma_i \mid \cdots] \sim Ber\left(\frac{p_i N(\beta_i \mid 0, c_i^2\tau_i^2)}{p_i N(\beta_i \mid 0, c_i^2\tau_i^2) + (1-p_i)N(\beta_i \mid 0, \tau_i^2)}\right),$$

and the remaining parameters are sampled simultaneously using an adaptive Metropolis algorithm [49] with multivariate normal proposals with an adaptive variance–covariance matrix. Additional details on the algorithm are given in Figure 2.

Final Model

As previously explained, the SSVS method can be used to reduce the set of possible models for further consideration, but in order to make inference or prediction on the annual maximum temperature, a final step has to be implemented. We opt for selecting the best model among those provided by the GEV-SSVS method applying a standard model selection criterion. First, we identify the most likely models according to the posterior distribution of the vector γ, and then we select the model with the lowest deviance information criterion (DIC) [50]. To fit these reduced models we simply set $p_i = 1$ and give a diffuse prior to the coefficients associated with the covariates by setting $c_i\tau_i = 10$.

Once this final model is selected, the vector of parameters of the model is reduced to $\boldsymbol{\theta} = (\beta_0, \sigma, \xi, \boldsymbol{\beta}^*)^\top$ where $\boldsymbol{\beta}^*$ is the vector of the coefficients of the covariates in the final selected model.

Following Gelman et al. [50], the DIC is a hierarchical modeling generalization of the Akaike information criterion (AIC), particularly useful in Bayesian model selection problems. It is defined as

$$DIC = -2\log[\mathbf{Y} \mid \hat{\boldsymbol{\theta}}_{Bayes}] + 2p_{DIC},$$

where $\boldsymbol{\theta} = (\beta_0, \sigma, \xi, \boldsymbol{\beta}^*)^\top$, $\hat{\boldsymbol{\theta}}_{Bayes}$ is the posterior mean $\mathrm{E}(\boldsymbol{\theta} \mid \mathbf{Y})$, and

$$p_{DIC} = 2\big(\log[\mathbf{Y} \mid \hat{\boldsymbol{\theta}}_{Bayes}] - \mathrm{E}(\log[\mathbf{Y} \mid \boldsymbol{\theta}])\big).$$

The idea is that models with smaller DIC should be preferred to models with larger DIC. Models are penalized both by the value of $-2\log[\mathbf{Y} \mid \hat{\boldsymbol{\theta}}_{Bayes}]$, which favors a good fit, but also by the effective number of parameters p_{DIC}.

Prediction and Inference

The final fitted Bayesian model can be used to obtain the posterior predictive distribution of the annual maximum temperature for a given vector of covariates, denoted by $[Y_t^{rep} \mid \mathbf{Y}]$, just integrating over the posterior distribution of parameters obtained from model fitting,

$$[Y_t^{rep} \mid \mathbf{Y}] = \int [Y_t^{rep} \mid \boldsymbol{\theta}][\boldsymbol{\theta} \mid \mathbf{Y}] d\boldsymbol{\theta}.$$

Samples of annual maximum temperatures with this posterior predictive distribution can be easily obtained from posterior samples of the parameters. As previously stated, the possibility of implementing this type of prediction to obtain replicates of the annual maximum series provides an important tool to make inference on different types of features of these series, for example record events. Using the simulated series, one can estimate not only any measure of interest but also the uncertainty of that estimation.

One of the most useful measures in the analysis of climate extremes is the return level z_p. Following Coles [15], this can be easily obtained from the inverse of the GEV distribution function as $z_p = G^{-1}(1-p)$ with G in (1). The return level z_p is the value that is exceeded with probability p, and it is associated with a return period of $1/p$ time units.

3. Results

This section presents a simulation study to assess the performance of the proposed algorithm. Then, an exploratory analysis of daily temperatures is performed and GEV models with a time-trend are fitted in each station to evaluate the change in the location and scale parameters. Second, a simple example of using the GEV-SSVS method to select a model for Barcelona based on surface temperature only is shown. Third, the proposed Bayesian variable selection method is applied to select a GEV model for projection. Finally, the selected model is validated, we show that the data do not deviate from the fitted distribution and that the model adequately captures the observed behavior. The main results are shown in detail for Barcelona and we refer the reader to the Appendix for the details of Madrid and Zaragoza.

3.1. Simulation Study

We propose a series of simulation experiments to assess the performance of the proposed GEV-SSVS algorithm. We focus on the case in which data $\mathbf{Y}$ are drawn from (2) with μ_t in (3) and the length of the series is $T = 100$. We set $\beta_0 = 30$ as a close reference to the temperature data. We study three values for the shape parameter $\xi = -0.3, 0, 0.2$, i.e., one for each distribution contained in the GEV. We set the variance of the data to 1 in all cases, leading to $\sigma = 1.01, 0.78, 0.55$, respectively. We consider two number of covariates $k = 10, 50$ and a vector of coefficients $\boldsymbol{\beta} = (1, 1, 0, \ldots, 0, 1, 1)^\top$, i.e., all zeros except the

first two and last two covariates. Finally, we study the performance of the algorithm with independent covariates and dependent covariates as follows.

For the independent covariates we independently generate for each covariate 100 independent standard normal data. Now we scale the covariates to have zero mean and unit variance before generating the **Y**'s. The dependent covariates are generated with a similar dependence to Yu et al. [36]. This involves the generation of k independent covariates as above. Then, we independently generate 100 additional standard normal data **Z** and add 2**Z** to each covariate. This results in a correlation between covariates of around 0.8. Now we scale the covariates to have zero mean and unit variance.

Each model is simulated 50 times and the results are summarized in Table 1. If $k = 10$ then $p_0 = 0.5$ and if $k = 50$ then $p_0 = 0.1$; the other parameters are left by default. The results show that for all models the true model is the most frequent in samples of γ at least 88% of the time, and is almost always among the promising models. Although the results seem similar between independent and dependent covariates, there is unsurprisingly an overall better performance with independent covariates than with dependent covariates.

Table 1. Summary from the 50 simulations with data generated from the 12 models in the three first columns. The first position is the proportion of simulations where the true model is the most frequent in γ. The 1st–10th position is the same but is within the 10 most frequent models, i.e., promising models. $\bar{x}$ and $q_{1/2}$ frequency are the mean and median across simulations of the frequency of appearance of the true model in γ.

	ξ	k	1st Position	1st–10th Position	$\bar{x}$ Frequency	$q_{1/2}$ Frequency
Independent	−0.3	10	1.00	1.00	0.69	0.71
	0	10	0.98	1.00	0.74	0.76
	0.2	10	1.00	1.00	0.84	0.86
	−0.3	50	0.92	1.00	0.61	0.67
	0	50	0.92	1.00	0.68	0.72
	0.2	50	1.00	1.00	0.82	0.84
Dependent	−0.3	10	0.88	1.00	0.41	0.42
	0	10	0.88	0.98	0.50	0.54
	0.2	10	0.94	1.00	0.63	0.70
	−0.3	50	0.90	1.00	0.51	0.54
	0	50	0.92	1.00	0.57	0.62
	0.2	50	0.98	1.00	0.64	0.68

3.2. Exploratory Data Analysis

3.2.1. Descriptive Analysis

This section shows a descriptive analysis of means, standard deviations, and linear trends in different versions of the temperature data introduced in Section 2.1. In particular, Table 2 shows some of these measurements for some subsets of the daily maximum temperature data from 1914, 1920, and 1951 up to year 2010 in Barcelona, Madrid, and Zaragoza. It considers daily maximum temperatures using all days in the year; using days in June, July and August (JJA); using days in July; or using the annual maximum temperature.

Table 2. Left: mean $\bar{x}$ and standard deviation s_x of daily maximum temperatures in JJA and July. Right: linear trend (°C/decade) in daily series, daily series of JJA, and annual maximum temperatures.

	JJA		July		Daily	JJA	max
Station	$\bar{x}$	s_x	$\bar{x}$	s_x		**Linear Trend**	
Barcelona	26.6	3.6	27.5	3.1	0.21	0.22	0.20
Madrid	29.5	4.3	31.0	3.5	0.13	0.18	0.16
Zaragoza	30.3	4.5	31.6	4.1	0.31	0.51	0.39

The daily maximum temperatures in JJA in Barcelona, Madrid. and Zaragoza are 26.6, 29.5 and 30.3 °C, respectively. The corresponding standard deviation is around 4 °C in the three series. The mean in July is between 1 and 1.5 °C higher and the standard deviation around 0.5 °C lower than in the three months JJA. To explore the evolution over time of temperature, we estimate the linear trend using ordinary least squares in three series of temperature: daily, daily in JJA and annual maximum. The trend in JJA ranges from 0.18 in Madrid to 0.51 °C/decade in Zaragoza. The trend in annual maximum temperature is lower in the three locations. The trend in the daily series using all days in the year is much lower in Madrid and Zaragoza than in the other two series. This different behavior confirms the need of a specific analysis to study the effect of climate change in temperature extremes. Figure 3 shows the evolution across years of the annual maximum temperature series. An increase of the level is clearly observed in the three series. This increase leads to a high number of observed records in the series.

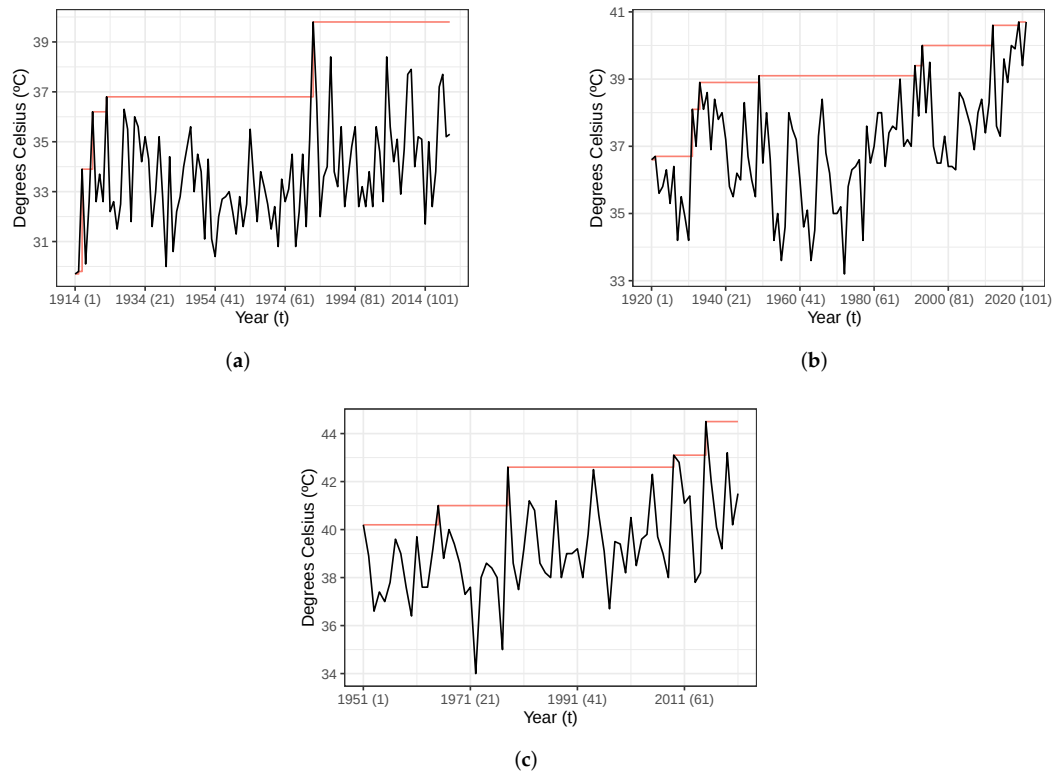

Figure 3. Annual maximum temperature series and their records. (**a**) Barcelona (1914–2021). (**b**) Madrid (1920–2021). (**c**) Zaragoza (1951–2021).

3.2.2. GEV Regression including Time-Trends

In order to analyze the need of a non-stationary GEV distribution to model the annual maximum temperature series, different models, including constant and time-varying location and scale parameters, are fitted. The estimation is carried out using a Bayesian approach and the R package **extRemes** [51] with the priors by default. These models could be fitted, except for the trend in scale, using the function available in **GEVSSVS** by setting $p_i = 1$. In fact, we find that our sampling algorithm mixes better and is more than 15 times faster when fitting these models for the same number of iterations. This difference increases with the number of covariates. However, we use **extRemes** in this section to make a fair comparison between the models that are not stationary in location and scale and to compare the results of this usual approach with our subsequent proposal. Note that we do not use the DIC provided by **extRemes**, we implement the expression explained above.

Table 3 summarizes the DIC of five GEV models: the first considers constant parameters μ and σ; the second includes a linear trend in location and a constant scale parameter;

the third is the same but with a quadratic trend in location; the fourth a linear trend in the scale parameter, $\log \sigma_t = \phi_0 + \phi_1 t$ and a constant location parameter, and the fifth a quadratic trend in scale. The lowest DIC is achieved by the model including a linear trend in location. The model with constant parameters and the model with a linear trend in scale yield a similar fit. According to these results, there is no evidence to include a time-varying scale parameter in the model. However, the decrease of DIC between the stationary GEV model and the models with a time-varying location suggests the need of including a time-varying location parameter. More details can be found in Table A1 of Appendix A, where the posterior distribution of the model parameters is summarized.

Table 3. DIC of GEV models including constant and time-varying location and scale parameters.

Model	Barcelona	Madrid	Zaragoza
i.i.d.	410	331	246
Linear trend in location	400	324	238
Quadratic trend in location	407	329	244
Linear trend in scale	411	331	243
Quadratic trend in scale	413	337	252

3.3. Modeling Annual Maximum Temperature

3.3.1. An Illustration of GEV-SSVS

In order to make clear the way of working with the variable selection method, we present the steps of its application in a simple example, the modeling of the annual maximum temperatures in Barcelona, but considering only a set of four potential covariates. The considered variables are the 2 m surface temperature $t2m$ in the four points that define the spatial grid where Barcelona is located (see Figure 1).

A preliminary step before applying the SSVS method in each dataset is to determine the values of p_i, c_i, and τ_i. The selected values here are $p_i = 1/2$, and the other parameters remain as in Section 2.2.2. Note that the value of p_i is related with the prior proportion of covariates we expect to keep in the model.

Table 4 summarizes the results of the application of the GEV-SSVS algorithm. The first column shows the four covariates, and the following ones the three models with a frequency higher than 5% in the MCMC samples. The rows below show the frequency and DIC of the models fitted for those covariates. The model with highest frequency corresponds to the model with variable $t2m$ in the SE point of the considered grid. Indeed, this is the model with the lowest DIC from all the 2^4 possible models. These results lead to select as final model the GEV distribution with location parameter including the covariate $t2m$ in the SE point.

Table 4. Values of γ with a frequency higher than 5% in the MCMC. Bottom: frequency of appearance and DIC. Barcelona (1914–2010).

Variable	Values of γ		
$t2m$ 2E–41N	0	1	1
$t2m$ 3E–41N	1	0	1
$t2m$ 2E–42N	0	0	0
$t2m$ 3E–42N	0	0	0
Frequency (%)	38	20	15
DIC	376	378	378

A summary of the posterior distribution of the model parameters for the selected model can be found in Table A2 of Appendix A.

3.3.2. GEV-SSVS with Atmospheric Covariates

Statistical downscaling methods that use ESM variables as input information to project require the use of variables that are adequately reproduced at the considered temporal and spatial scale. This condition prevents the use in our model of surface temperature as input variables, since they are in general badly reproduced at a daily and local scale. However, variables measured at a given elevation are less influenced by the relief of the area, better reproduced by ESM and, consequently, they can be used more safely as input. In addition, we should take into account that information on atmospheric conditions (not only in the considered location, but also in an area around it) may provide useful information to model annual maximum temperature. Thus, for a temperature series in a given location, we will consider the $4 \times 3 \times 4 = 48$ variables corresponding to the four grid points around the location, three variables (temperature, specific humidity, and geopotential) and four pressure levels (850, 700, 500, and 300 hPa). All covariates are scaled in mean and variance for the implementation of the algorithm. It is also noteworthy that there is a high correlation between the same variable measured at the four points of the grid and also between the different variables measured at the same point. This high correlation causes different covariates to produce a similar goodness of fit, returning very similar frequencies of appearance in the selection algorithm. For this reason we suggest to run several chains (we run 5) until consistent results are obtained between them, and then compare the best returned models by means of their DIC.

The final model for Barcelona, according to the GEV-SVSS approach described in Section 2.2.2, includes the variable *t.850* at 2E–41N and 3E–42N, that is, two temperatures at 850 hPa pressure level at the SW and NE points. The posterior mean and credible interval (CI) of the model parameters are summarized in Table 5. Both covariates have a very high correlation of 0.98, but both are significant, with the posterior mean of the SW point coefficient being negative while that of the NE is positive, with the latter being of greater magnitude. This means that the annual maximum increases with the increase in temperature *t.850* at the NE point located in the interior of Iberian Peninsula, while it is regulated by the same variable at the SW point located in the Mediterranean Sea. The 95% CI of the shape parameter ξ contains zero, consequently the posterior distribution could be considered a Gumbel distribution. The model is satisfactorily validated according to usual diagnostic plots, see Figure A1 in Appendix B.

As a product of the fitted model, the left plot in Figure 4 shows the posterior mean of the 10 and 20 year return values obtained from the model, together with the observed annual maximum temperature as a reference. As the observed series, the posterior mean of the return values shows a non-monotonous increasing evolution. The right plot in Figure 4 shows the posterior probability of the annual maximum temperature being higher than 36, 37, and 38 °C. This signal also shows a non-monotonous increasing evolution, specially from 1980 onwards.

Table 5. Posterior mean and 95% CI of the final GEV model parameters according to the GEV-SSVS selection. Barcelona (1914–2010).

$DIC = 383$	Mean	CI
β_0	32.73	$(32.40, 33.07)$
t.850 2E–41N	-1.74	$(-3.36, -0.02)$
t.850 3E–42N	2.55	$(0.84, 4.17)$
σ	1.51	$(1.28, 1.77)$
ξ	-0.05	$(-0.18, 0.10)$

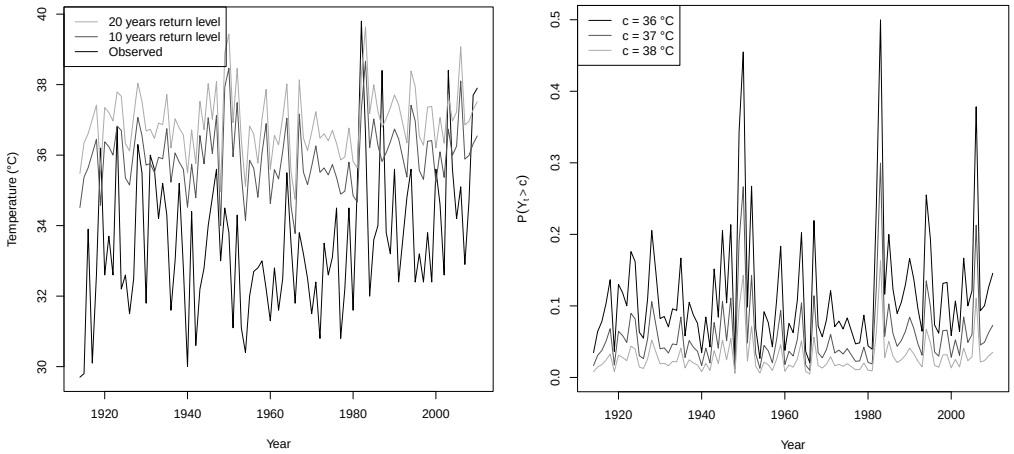

Figure 4. Left: posterior mean of 10 and 20 year return values by year. **Right**: posterior probability of the annual maximum temperature being higher than 36, 37, and 38 °C by year according to the final GEV model. Barcelona (1914–2010).

The final model for Madrid is very similar to the model for Barcelona; both have two temperatures at 850 hPa *t.850* as covariates. The summary of the model parameters are in Table A3 of Appendix B; the main difference is that ξ is significantly negative, indicating a Weibull distribution. Diagnostic plots are satisfactory, as shown in Figure A2 of Appendix B.

The final model for Zaragoza is different from the others. It includes the covariates *q.850* 1W–42N and *z.300* 0E–41N, that is, the specific humidity at 850 hPa in the NW point and the geopotential at 300 hPa in the SE point. The summary of the model parameters are in Table A4 of Appendix B. The two covariates are not significantly correlated. The posterior mean value of the geopotential coefficient is positive, while that of humidity is negative. This means that a situation with low humidity at 850 hPa and high geopotential at 300 hPa height, features associated to a persistent anticyclone situation over the North Atlantic, increases the risk of high annual maximum temperature. The obtained shape parameter is compatible with a Weibull distribution. Diagnostic plots in Figure A3 in Appendix B are satisfactory and show a good fit.

All the models include information from different grid points, even if this information is highly correlated and comes from the same type of measurement. This confirms the convenience of considering as possible covariates at least the variables corresponding to the four points of the spatial grid where the station is located.

It is noteworthy that the models obtained from applying the GEV-SSVS approach clearly improve the DIC of the models fitted in Section 3.2.2 with a polynomial trend in time in the location parameter. This is, a DIC of 383 vs. 400, 315 vs. 324, and 222 vs. 238 in Barcelona, Madrid, and Zaragoza, respectively.

4. Conclusions and Future Work

This work proposes a new Bayesian approach for variable selection in a GEV regression framework. The approach consists of two steps: the first one is a GEV-SSVS method and aims to identify the sets of covariates that lead to the models with the highest posterior probabilities. In some cases, for example with sets of highly correlated variables, the previous approach may yield a large number of models with similar probabilities. Consequently, a second step to select the best model is implemented by applying a standard model selection criterion, the DIC. In cases of serious problems of multicollinearity, it is convenient to implement the GEV-SSVS approach by simulating a high number of iterations using several chains, and to check that the results are similar in all of them.

The suggested GEV-SSVS is implemented in the R package **GEVSSVS**, available at https://github.com/JorgeCastilloMateo/GEVSSVS (accessed on 10 December 2022). This package also allows the fit of a wide range of GEV models, including non-linear terms and interactions in the location parameter. The sampling algorithm implemented in the

functions in **GEVSSVS** mixes better and is much faster than the Bayesian algorithms used in the package **extRemes**. These differences increase with the number of covariates.

The proposed method is used to select the best atmospheric covariates to model the annual maximum temperature in three Spanish locations with different climatic conditions. The variable selection method is applied to a set of 48 potentially influential atmospheric covariates, which are highly correlated. In the three series, the approach leads to parsimonious models including two covariates and provides a clear interpretation of the predictor effects. The selected covariates correspond to different points in the grid surrounding the location where the series is measured, and at least one of the predictors corresponds to a 850 hPa pressure level.

A difficulty associated to variable selection in GEV regression appears in finding the subset that maximizes the likelihood, as is also the case for generalized linear and non-linear models where the full conditional densities cannot be obtained directly; the resulting mixture posterior may be difficult to sample using standard MCMC methods due to multimodality. Future work will consider specific MCMC algorithms that automatically tune the parameters of a family of mixture proposal distributions during simulation [52]. This approach is necessary to implement the variable selection method when a high number of potential covariates is available. Another research line is to generalize the procedure to a spatio-temporal framework. In that case the selection method should aim to select the set of relevant covariates that could be different in each location.

From a climate point of view, an interesting application will be the use of the proposed approach to select and fit models that can be used to downscale the outputs of ESM under climate change scenarios. In particular, the fitted model in this work could be used to obtain downscaled projections of annual maximum temperatures for the next 50 years.

Author Contributions: Conceptualization, J.C.-M. and J.A. (Jesús Asín); methodology, J.C.-M. and J.A. (Jesús Asín); software, J.C.-M.; validation, J.C.-M., J.A. (Jesús Asín), A.C.C. and J.A. (Jesús Abaurrea); formal analysis, J.A. (Jesús Asín) and J.C.-M.; investigation, J.C.-M., J.A. (Jesús Asín) and A.C.C.; data curation, J.A. (Jesús Asín), J.C.-M. and J.M.-L.; writing—original draft preparation, J.C.-M., A.C.C. and J.A. (Jesús Asín); writing—review and editing, A.C.C., J.A. (Jesús Asín), J.C.-M., J.M.-L. and J.A. (Jesús Abaurrea); visualization, J.A. (Jesús Asín) and J.C.-M.; supervision, J.C.-M.; project administration, J.C.-M.; funding acquisition, J.M.-L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data (daily temperature) and metadata are provided by the ECA&D project and available at http://www.ecad.eu (accessed on 1 April 2022). The series of Barcelona is the blended series of station SPAIN, BARCELONA-FABRA OBSERVATORY (STAID: 335). The series of Madrid is the blended series of station SPAIN, MADRID-RETIRO (STAID: 230). The series of Zaragoza is the blended series of station SPAIN, ZARAGOZA AEROPUERTO (STAID: 238).

Acknowledgments: This work was partially supported by the Ministerio de Ciencia e Innovación under Grants PID2020-116873GB-I00 and TED2021-130702B-I00; Gobierno de Aragón under Research Group E46_20R: Modelos Estocásticos; and J.C.-M. was supported by Gobierno de Aragón under Doctoral Scholarship ORDEN CUS/581/2020. The authors thank the ECA&D project and the ECMWF for providing the data. The authors are grateful to the Editor and two Referees for their insightful and constructive remarks on an earlier version of the paper.

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

Appendix A. Preliminary Models

Table A1 shows the posterior mean and 95% CI of the models described in Section 3.2.2. And Table A2 shows the same for the model selected in Section 3.3.1.

Table A1. Posterior mean and 95% CI of the GEV models including constant parameters and linear time-trends in location or scale.

		Barcelona		Madrid		Zaragoza	
		Mean	**CI**	**Mean**	**CI**	**Mean**	**CI**
i.i.d.	μ	32.67	(32.21, 33.13)	36.22	(35.81, 36.61)	38.30	(37.70, 38.87)
	σ	1.80	(1.51, 2.17)	1.53	(1.28, 1.83)	1.83	(1.48, 2.32)
	ξ	−0.10	(−0.23, 0.07)	−0.32	(−0.45, −0.18)	−0.25	(−0.41, −0.08)
t–location	β_0	31.49	(30.60, 32.39)	35.52	(34.84, 36.18)	37.15	(36.14, 38.14)
	β_1	0.02	(0.01, 0.04)	0.02	(0.00, 0.03)	0.04	(0.01, 0.07)
	σ	1.63	(1.34, 1.99)	1.48	(1.25, 1.78)	1.73	(1.41, 2.17)
	ξ	−0.03	(−0.21, 0.18)	−0.34	(−0.49, −0.19)	−0.28	(−0.45, −0.11)
t–scale	μ	32.71	(32.26, 33.17)	36.20	(35.81, 36.57)	38.18	(37.60, 38.76)
	ϕ_0	0.68	(0.36, 1.05)	0.30	(0.02, 0.63)	0.32	(−0.05, 0.74)
	ϕ_1	0.00	(−0.01, 0.00)	0.00	(0.00, 0.01)	0.01	(0.00, 0.02)
	ξ	−0.05	(−0.24, 0.17)	−0.36	(−0.51, −0.20)	−0.30	(−0.46, −0.12)

Table A2. Posterior mean and 95% CI of the GEV model parameters according to the GEV-SSVS selection between the 2 m surface temperature *t2m* variables. Barcelona (1914–2010).

$DIC = 376$	**Mean**	**CI**
β_0	32.79	(32.46, 33.13)
t2m 3E–41N	1.01	(0.69, 1.33)
σ	1.50	(1.28, 1.77)
ξ	−0.10	(−0.22, 0.05)

Appendix B. Results of the GEV Models According to the GEV-SSVS Selection

Figures A1, A2 and A3 show the usual diagnostic plots for the final GEV models according to the GEV-SSVS selection in Barcelona, Madrid and Zaragoza, respectively. Summaries for the posterior distribution of the model parameters in the final models for the last two stations are shown in Tables A3 and A4.

Table A3. Posterior mean and 95% CI of the final GEV model parameters according to the GEV-SSVS selection. Madrid (1920–2010).

$DIC = 315$	**Mean**	**CI**
β_0	36.30	(35.97, 36.62)
t.850 4W–40N	2.13	(0.80, 3.34)
t.850 4W–41N	−1.64	(−2.76, −0.35)
σ	1.44	(1.22, 1.72)
ξ	−0.39	(−0.55, −0.24)

Table A4. Posterior mean and 95% CI of the final GEV model parameters according to the GEV-SSVS selection. Zaragoza (1951–2010).

$DIC = 222$	**Mean**	**CI**
β_0	38.34	(37.95, 38.73)
q.850 1W–42N	−0.55	(−0.93, −0.17)
z.300 0E–41N	0.92	(0.54, 1.29)
σ	1.38	(1.13, 1.70)
ξ	−0.13	(−0.29, 0.04)

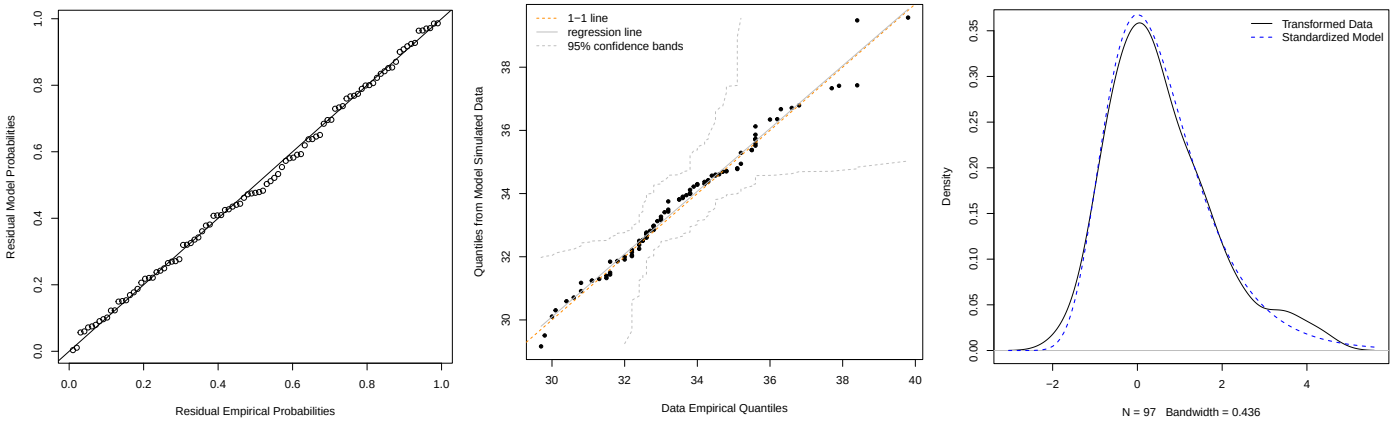

Figure A1. Diagnostic plots of the final GEV model parameters according to the GEV-SSVS selection. Barcelona (1914–2010).

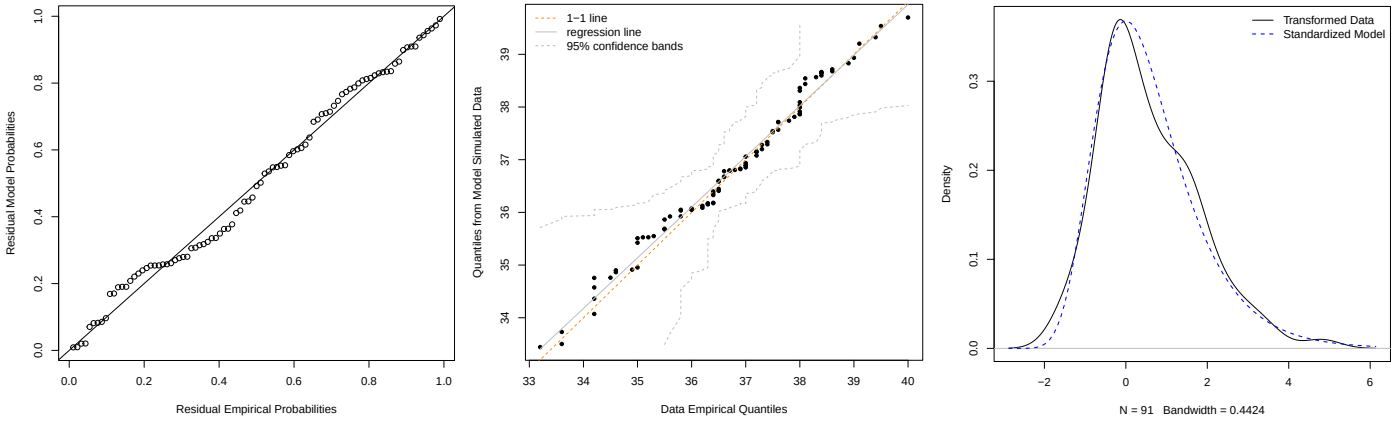

Figure A2. Diagnostic plots of the final GEV model parameters according to the GEV-SSVS selection. Madrid (1920–2010).

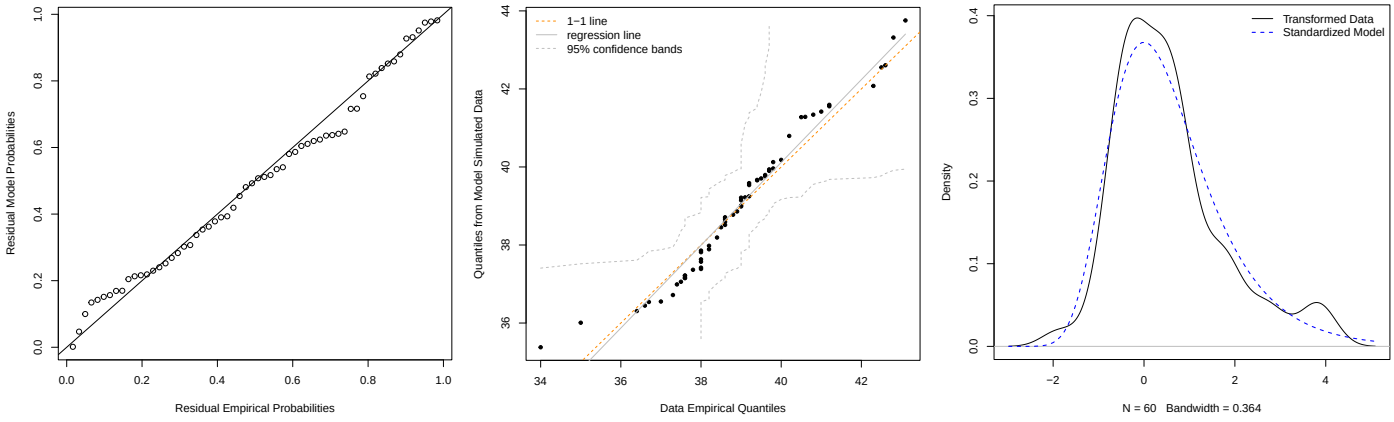

Figure A3. Diagnostic plots of the final GEV model parameters according to the GEV-SSVS selection. Zaragoza (1951–2010).

References

1. Rahmstorf, S.; Coumou, D. Increase of extreme events in a warming world. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17905–17909. [CrossRef]
2. IPCC. Summary for Policymakers. In *Global Warming of 1.5 °C. An IPCC Special Report on the Impacts of Global Warming of 1.5 °C above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways*; Masson-Delmotte, V., Zhai, P., Pörtner, H.O., Roberts, D., Skea, J., Shukla, P.R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., et al., Eds.; World Meteorological Organization: Geneva, Switzerland, 2018.

3. Castillo-Mateo, J.; Lafuente, M.; Asín, J.; Cebrián, A.C.; Gelfand, A.E.; Abaurrea, J. Spatial modeling of day-within-year temperature time series: An examination of daily maximum temperatures in Aragón, Spain. *J. Agric. Biol. Environ. Stat.* **2022**, *27*, 487–505. [CrossRef]
4. Castillo-Mateo, J.; Asín, J.; Cebrián, A.C.; Gelfand, A.E.; Abaurrea, J. Spatial quantile autoregression for season within year daily maximum temperature data. *Ann. Appl. Stat. in press*. [CrossRef]
5. Amengual, A.; Homar, V.; Romero, R.; Brooks, H.; Ramis, C.; Gordaliza, M.; Alonso, S. Projections of heat waves with high impact on human health in Europe. *Glob. Planet. Chang.* **2014**, *119*, 71–84. [CrossRef]
6. Tobías, A.; Armstrong, B.; Gasparrini, A.; Diaz, J. Effects of high summer temperatures on mortality in 50 Spanish cities. *Environ. Health* **2014**, *13*, 48. [CrossRef] [PubMed]
7. Parmesan, C.; Root, T.L.; Willig, M.R. Impacts of extreme weather and climate on terrestrial biota. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 443–450. [CrossRef]
8. Sang, H.; Gelfand, A.E. Hierarchical modeling for extreme values observed over space and time. *Environ. Ecol. Stat.* **2009**, *16*, 407–426. [CrossRef]
9. Goubanova, K.; Li, L. Extremes in temperature and precipitation around the Mediterranean basin in an ensemble of future climate scenario simulations. *Glob. Planet. Chang.* **2007**, *57*, 27–42. [CrossRef]
10. Kharin, V.; Zwiers, F.; Zhang, X.; Wehner, M. Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Clim. Chang.* **2013**, *119*, 345–357. [CrossRef]
11. Gao, M.; Zheng, H. Nonstationary extreme value analysis of temperature extremes in China. *Stoch. Environ. Res. Risk Assess.* **2018**, *32*, 1299–1315. [CrossRef]
12. Paola, F.; Giugni, M.; Pugliese, F.; Annis, A.; Nardi, F. GEV parameter estimation and stationary vs. non-Stationary analysis of extreme rainfall in African test cities. *Hydrology* **2018**, *5*, 28. [CrossRef]
13. Stein, M. Should annual maximum temperatures follow a generalized extreme value distribution? *Biometrika* **2017**, *104*, 1–16. [CrossRef]
14. Cooley, D. Extreme value analysis and the study of climate change: A commentary on Wigley 1988. *Clim. Chang.* **2009**, *97*, 77–83. [CrossRef]
15. Coles, S. *An Introduction to Statistical Modeling of Extreme Values*; Springer: Berlin/Heidelberg, Germany, 2001.
16. Smith, R.L. Likelihood and modified likelihood estimation for distributions with unknown endpoints. In *Recent Advances in Life-Testing and Reliability*, 1st ed.; Balakrishnan, N., Ed.; CRC Press: Boca Raton, FL, USA, 1995; pp. 455–474.
17. Aguirre-Salado, A.I.; Aguirre-Salado, C.A.; Alvarado, E.; Santiago-Santos, A.; Lancho-Romero, G.A. On the smoothing of the generalized extreme value distribution parameters using penalized maximum likelihood: A case study on UVB radiation maxima in the Mexico City Metropolitan Area. *Mathematics* **2020**, *8*, 329. [CrossRef]
18. Gaetan, C.; Grigoletto, M. A hierarchical model for the analysis of spatial rainfall extremes. *J. Agric. Biol. Environ. Stat.* **2007**, *12*, 434. [CrossRef]
19. Reich, B.; Shaby, B.; Cooley, D. A hierarchical model for serially-dependent extremes: A study of heat waves in the Western US. *J. Agric. Biol. Environ. Stat.* **2014**, *19*, 119–135. [CrossRef]
20. Friederichs, P.; Thorarinsdottir, T.L. Forecast verification for extreme value distributions with an application to probabilistic peak wind prediction. *Environmetrics* **2012**, *23*, 579–594. [CrossRef]
21. Abaurrea, J.; Asín, J. Forecasting local daily precipitation patterns in a climate change scenario. *Clim. Res.* **2005**, *28*, 183–197. [CrossRef]
22. Fischer, E.M.; Sippel, S.; Knutti, R. Increasing probability of record-shattering climate extremes. *Nat. Clim. Chang.* **2021**, *11*, 689–695. [CrossRef]
23. Wehrli, K.; Luo, F.; Hauser, M.; Shiogama, H.; Tokuda, D.; Kim, H.; Coumou, D.; May, W.; Le Sager, P.; Selten, F.; et al. The ExtremeX global climate model experiment: Investigating thermodynamic and dynamic processes contributing to weather and climate extremes. *Earth Syst. Dyn.* **2022**, *13*, 1167–1196. [CrossRef]
24. Abaurrea, J.; Asín, J.; Cebrián, A.C. Modeling and projecting the occurrence of bivariate extreme heat events using a non-homogeneous common Poisson shock process. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 309–322. [CrossRef]
25. Abaurrea, J.; Asín, J.; Cebrián, A. Modelling the occurrence of heat waves in maximum and minimum temperatures over Spain and projections for the period 2031-60. *Glob. Planet. Chang.* **2018**, *161*, 244–260. [CrossRef]
26. Keellings, D.; Waylen, P. Investigating teleconnection drivers of bivariate heat waves in Florida using extreme value analysis. *Clim. Dyn.* **2015**, *44*, 3383–3391. [CrossRef]
27. Van Oldenborgh, G.J.; Wehner, M.F.; Vautard, R.; Otto, F.E.; Seneviratne, S.I.; Stott, P.A.; Hegerl, G.C.; Philip, S.Y.; Kew, S.F. Attributing and projecting heatwaves is hard: We can do better. *Earth's Future* **2022**, *10*, e2021EF002271. [CrossRef]
28. Seong, M.G.; Min, S.K.; Zhang, X. A Bayesian attribution analysis of extreme temperature changes at global and regional scales. *J. Clim.* **2022**, *35*, 4589–4603. [CrossRef]
29. Fernandez-Granja, J.A.; Casanueva, A.; Bedia, J.; Fernandez, J. Improved atmospheric circulation over Europe by the new generation of CMIP6 earth system models. *Clim. Dyn.* **2021**, *56*, 3527–3540. [CrossRef]
30. Iturbide, M.; Casanueva, A.; Bedia, J.; Herrera, S.; Milovac, J.; Gutiérrez, J.M. On the need of bias adjustment for more plausible climate change projections of extreme heat. *Atmos. Sci. Lett.* **2022**, *23*, e1072. [CrossRef]
31. Miller, A.J. *Subset Selection in Regression*, 1st ed.; Chapman and Hall: New York, NY, USA, 1990.

32. Gramacy, R.B. *Surrogates: Gaussian Process Modeling, Design, and Optimization for the Applied Sciences*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2020.
33. George, E.I.; McCulloch, R.E. Variable selection via Gibbs sampling. *J. Am. Stat. Assoc.* **1993**, *88*, 881–889. [CrossRef]
34. Scheipl, F. **spikeSlabGAM**: Bayesian variable selection, model choice and regularization for generalized additive mixed models in R. *J. Stat. Softw.* **2011**, *43*, 1–24. [CrossRef]
35. He, V.X.; Wand, M.P. Bayesian generalized additive model selection including a fast variational option. *arXiv* **2022**, arXiv:2201.00412.
36. Yu, K.; Chen, C.W.S.; Reed, C.; Dunson, D.B. Bayesian variable selection in quantile regression. *Clim. Dyn.* **2013**, *6*, 261–274. [CrossRef]
37. El Adlouni, S.; Ouarda, T. Joint Bayesian model selection and parameter estimation of the generalized extreme value model with covariates using birth-death Markov chain Monte Carlo. *Water Resour. Res.* **2009**, *45*, W06403. [CrossRef]
38. Cebrián, A.C.; Castillo-Mateo, J.; Asín, J. Record tests to detect non-stationarity in the tails with an application to climate change. *Stoch. Environ. Res. Risk Assess.* **2022**, *36*, 313–330. [CrossRef]
39. Castillo-Mateo, J. Distribution-free changepoint detection tests based on the breaking of records. *Environ. Ecol. Stat.* **2022**, *29*, 655–676. [CrossRef]
40. Yazdi, M.; Kabir, S.; Walker, M. Uncertainty handling in fault tree based risk assessment: State of the art and future perspectives. *Process. Saf. Environ. Prot.* **2019**, *131*, 89–104. [CrossRef]
41. Clyde, M.; George, E.I. Model Uncertainty. *Stat. Sci.* **2004**, *19*, 81–94. [CrossRef]
42. Perrakis, K.; Ntzoufras, I. Stochastic Search Variable Selection (SSVS). In *Wiley StatsRef: Statistics Reference Online*; Wiley: Hoboken, NJ, USA, 2015; pp. 1–6. [CrossRef]
43. Klein Tank, A.M.G.; Wijngaard, J.B.; Können, G.P.; Böhm, R.; Demarée, G.; Gocheva, A.; Mileta, M.; Pashiardis, S.; Hejkrlik, L.; Kern-Hansen, C.; et al. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *Int. J. Climatol.* **2002**, *22*, 1441–1453. [CrossRef]
44. Poli, P.; Hersbach, H.; Dee, D.P.; Berrisford, P.; Simmons, A.J.; Vitart, F.; Laloyaux, P.; Tan, D.G.; Peubey, C.; Thépaut, J.N.; et al. ERA-20C: An atmospheric reanalysis of the twentieth century. *J. Clim.* **2016**, *29*, 4083–4097. [CrossRef]
45. Hufkens, K.; Stauffer, R.; Campitelli, E. *The **ecwmfr** Package: An Interface to ECMWF API Endpoints*; Zenodo: Geneva, Switzerland, 2019. [CrossRef]
46. Fisher, R.A.; Tippett, L.H.C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Camb. Philos. Soc.* **1928**, *24*, 180–190. [CrossRef]
47. Leadbetter, M.R.; Lindgren, G.; Rootzen, H. *Extremes and Related Properties of Random Sequences and Processes*, 1st ed.; Springer: New York, NY, USA, 1983. [CrossRef]
48. Gilks, W.R.; Richardson, S.; Spiegelhalter, D. *Markov Chain Monte Carlo in Practice*; Chapman and Hall/CRC: Boca Raton, FL, USA, 1995.
49. Haario, H.; Saksman, E.; Tamminen, J. An adaptive Metropolis algorithm. *Bernoulli* **2001**, *7*, 223–242. [CrossRef]
50. Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, D.B.; Vehtari, A.; Rubin, D.B. *Bayesian Data Analysis*, 3rd ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2013.
51. Gilleland, E.; Katz, R.W. **extRemes** 2.0: An extreme value analysis package in R. *J. Stat. Softw.* **2016**, *72*, 1–39. [CrossRef]
52. Ji, C.; Schmidler, S.C. Adaptive Markov Chain Monte Carlo for Bayesian Variable Selection. *J. Comput. Graph. Stat.* **2013**, *22*, 708–728. [CrossRef]