



Research papers

Nonstationary frequency analysis of extreme precipitation: Embracing trends in observations

Gabriel Anzolin ^{a,*}, Debora Y. de Oliveira ^b, Jasper A. Vrugt ^b, Amir Aghakouchak ^b, Pedro L. B. Chaffe ^c

^a Graduate Program in Environmental Engineering, Federal University of Santa Catarina, Florianopolis, Brazil

^b Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA

^c Department of Sanitary and Environmental Engineering, Federal University of Santa Catarina, Florianopolis, Brazil

ARTICLE INFO

Keywords:

Frequency Analysis
Nonstationarity
Climate Change
Bayesian Inference

ABSTRACT

Knowledge of the recurrence intervals of precipitation extremes is vital for infrastructure design, risk assessment, and insurance planning. However, trends and shifts in rainfall patterns globally pose challenges to the application of extreme value analysis (EVA) which relies critically on the assumption of stationarity. In this paper, we explore: (1) the suitability of nonstationary (NS) models in the presence of statistically significant trends, and (2) their potential in modeling out-of-sample data to improve frequency analysis of extreme precipitation. We analyze the benefits of using a nonstationary Generalized Extreme Value (GEV) model for annual extreme precipitation records from Southern Brazil. The location of the GEV distribution is allowed to change with time. The unknown GEV model parameters are estimated using Bayesian techniques coupled with Markov chain Monte Carlo simulation. Next, we use GAME sampling to compute the evidence (and their ratios, the so-called Bayes factors) for stationary and nonstationary models of annual maximum precipitation. Our results show that the presence of a statistically significant trend in annual maximum precipitation alone does not justify the use of a NS model. The location parameter of the GEV distribution must also be well defined, otherwise, stationary models of annual maximum precipitation receive more support by the data. These findings reiterate the importance of accounting for GEV model parameters and predictive uncertainty in frequency analysis and hypothesis testing of annual maximum precipitation data records. Furthermore, a meaningful EVA demands detailed knowledge about the origin and persistence of observed changes.

1. Introduction

Understanding the frequency and intensity of rainfall extremes is key to infrastructure design and risk assessment (Hailegeorgis and Alfredsen, 2017; Veneziano et al., 2006). Increases in Earth's temperature in recent decades (Barnett et al., 1999; Cheng et al., 2015), alongside with shifts in hydroclimatic and atmospheric circulation patterns (Milly et al., 2015, 2008), have altered the occurrence frequency of extreme rainfall (Barbero et al., 2019; Cheng and Aghakouchak, 2014; Moustakis et al., 2021; Papalexiou and Montanari, 2019; Prein and Mearns, 2021; Rago et al., 2019; Touma et al., 2022), and changed significantly its intensity, as outlined by IPCC reports (Pachauri et al., 2014). Therefore, understanding the trends and patterns of future extreme precipitation and developing quantitative methods for describing statistics of changing extremes is essential for society's resilience (Milly et al., 2015, 2008).

Typically, the estimation of design rainfall relies on Extreme Value Analysis (EVA), a prevalent approach that estimates the recurrence of rare events through probabilistic methods (Huard et al., 2010; Nerkantzaki and Papalexiou, 2019; Papalexiou and Koutsoyiannis, 2013). This involves techniques built on the assumption of stationarity, where extreme rainfall events are described as samples from some extreme value distribution with fixed parameters resulting in time-invariant statistics of extremes (Petrow and Merz, 2009; Read and Vogel, 2015; Serago and Vogel, 2018; Vogel et al., 2011). This approach ignores trends or changes in the precipitation distribution over time (Milly et al., 2008; Papalexiou et al., 2018).

One way to account for changes in extreme rainfall is by using the output of Global Climate Models (GCM) as input to EVA, enabling an assessment of both current and future extreme rainfall intensity and frequency (Bador et al., 2018; Emmanouil et al., 2023; Fadhel et al.,

* Corresponding author at: Department of Sanitary and Environmental Engineering, Federal University of Santa Catarina, Florianopolis, Brazil.
E-mail address: gabriel_anzolin@hotmail.com (G. Anzolin).

2017; Lopez-Cantu et al., 2020; Prein et al., 2017; Ragno et al., 2018). However, the output of GCMs is subject to considerable bias and uncertainty due to limitations and errors in process parametrization, insufficient spatial resolution, and inaccurate characterization of natural and anthropogenic climate variability (e.g., Bador et al., 2018; Cook et al., 2017, 2020; Emmanouil et al., 2023; Fadhel et al., 2017; Haughton et al., 2015; Lopez-Cantu et al., 2020; Ragno et al., 2018). This impacts the simulation accuracy of extreme precipitation, particularly in regions with complex terrain, convective rainfall, and strong seasonal variability (e.g., Barbero et al., 2019; Fatichi et al., 2016; Fischer et al., 2013; GregerSEN et al., 2013; Kendon et al., 2017; Mousatikis et al., 2020; Pereima et al., 2022).

To overcome those limitations, nonstationary (NS) models – which allow the mean, variance or shape of an extreme value distribution to change over time – are a pragmatic alternative to analyzing the frequency and recurrence interval of extreme rainfall. Yet, the potential usefulness of such nonstationary models of climatic variables is actively debated in the hydrologic literature and there are conflicting opinions on how and when such models should be applied (Cheng and Aghakouchak, 2014; Milly et al., 2008, 2015; Montanari and Koutsoyiannis, 2014; Ragno et al., 2019; Serinaldi and Kilsby, 2015). The use of a non-constant mean of the extreme value distribution increases the number of estimable model parameters, for example. This enlarges parameter uncertainty and if juxtaposed with epistemic uncertainty due to (among others) the use of inappropriate covariates (Emmanouil et al., 2022) requires a leap of faith in the application of nonstationary models. While nonstationary models enhance the description of recurrence intervals within-sample (e.g., Cheng and Aghakouchak, 2014; Ouarda et al.,

2020, 2019; Ragno et al., 2019; Serago and Vogel, 2018; Šraj et al., 2016; Vu and Mishra, 2019 and many others), fewer studies have demonstrated their superiority in predicting the frequency of out-of-sample extreme events (e.g., Anzolin et al., 2023; Lee et al., 2020; Luke et al., 2017; Roderick et al., 2020). There is still a need to assess whether nonstationary models can be used to predict the frequency of extreme rainfall, especially out-of-sample, thereby properly taking into account model parameters and predictive uncertainty.

Here, we test: (1) the suitability of NS models in the presence of statistically significant trends, and (2) their potential in modeling out-of-sample data to understand whether the presence of significant trends justifies the use of a nonstationary frequency analysis model. We analyze daily annual maximum rainfall records from Southern Brazil and evaluate the performance of the GEV distribution by comparing a constant and a non-constant location parameter (and hence, a constant and nonconstant mean). We implement a Bayesian method and draw inferences from the data through the posterior distribution of the GEV model parameters using Bayes factors. This approach explicitly accounts for model parameters and predictive uncertainty in hypothesis testing and model selection.

2. Material and methods

2.1. Study area and data

Southern Brazil (SB), is in the transition of tropical and extratropical climates (Fig. 1a), a sensitive area to a changing climate, susceptible to the expansion of the tropics in the Southern Hemisphere (Lucas et al.,

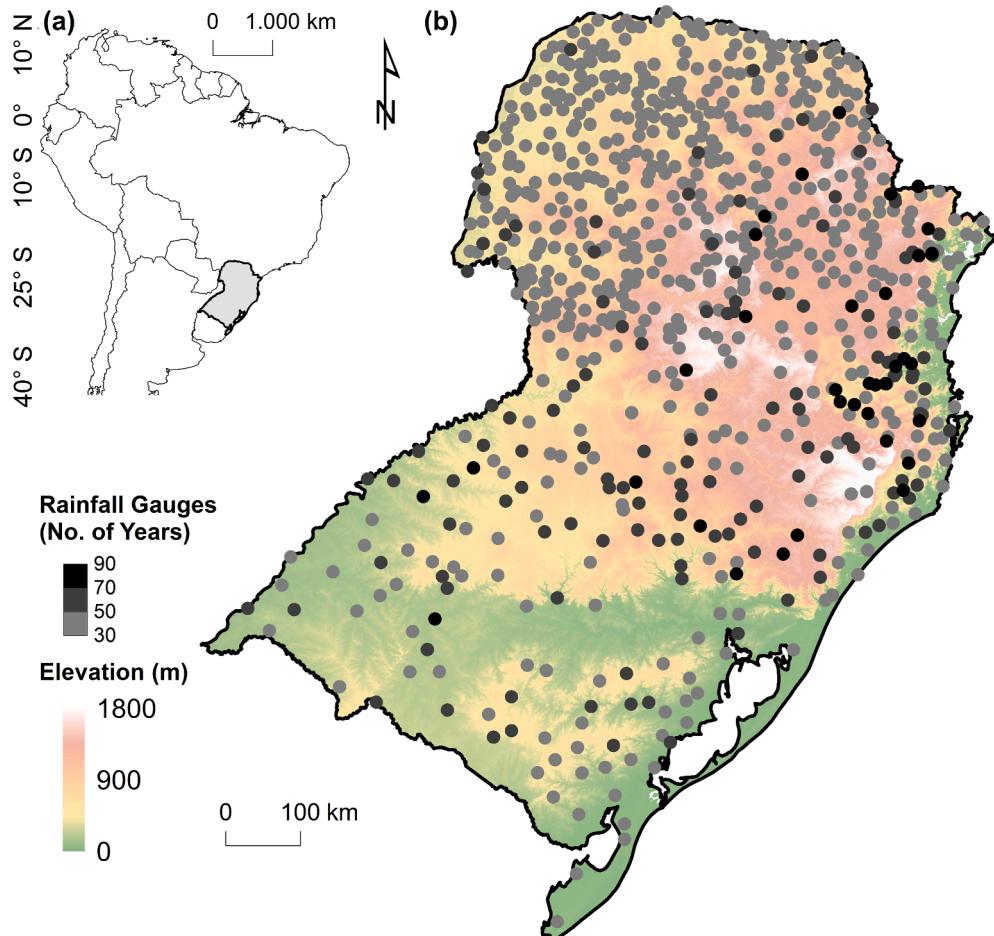


Fig. 1. (a) Location of South America, Brazil and Southern Brazil (SB), (b) surface elevation above the mean sea level (shading) for SB, and location of the 777 rainfall gauges selected for this study.

2014). The region has a complex topography that creates a diversity of rainfall generation mechanisms, with frontal rainfall in high areas, and a high occurrence of convective rainfall in the summer months. It is also characterized by varying spatial seasonality in precipitation, with summer monsoons in the north region and no seasonality in the remaining areas (Cavalcanti et al., 2009). This is especially relevant considering that GCM models do not perform well in this transition zone (Pereima et al., 2022).

We use daily rainfall data from the HidroWeb Portal of the Brazilian National Water Agency (ANA; <https://www.snh.gov.br/hidroweb/>). We screen gauges based on their available record length and data quality as in CAMELS-BR (Chagas et al., 2020). We select rainfall gauges with at least 30 years of data with less than 5 % of missing values. Gauges with spurious values (e.g., data with incorrect order of magnitudes, or zeros in place of missing data) are not included in the analysis. Fig. 1b shows the 777 high-quality rainfall gauges that passed our quality control procedure. For each rainfall gauge, we generate annual maximum series (AMS), by extracting the maximum value of each year. The full record of each AMS is a vector $\mathbf{P} = \{P_1, \dots, P_N\}$, where N is the number of available years. First, we define the fitting period data, $\mathbf{X} = \{P_1, \dots, P_n\}$, where n is the number of years in the fitting period. In the case of the within-sample analysis, the vector \mathbf{X} has n equal to the total number of observations – the full record is used as fitting period. For the out-of-sample analysis, we split the data into two periods: (1) a fitting period with the first 30 observations ($n = 30$), and (2) an evaluation period with the remaining observations, i.e., $\mathbf{X}^* = \{P_{n+1}, \dots, P_N\}$. For the out-sample-sample analysis, we only use AMS with at least 40 years of observations (358 records), to ensure 30 years of observations for the fitting period and at least 10 years to test the model performance.

2.2. Frequency analysis

We use the generalized extreme value (GEV) distribution in the extreme rainfall frequency analysis. This distribution summarizes three families of extreme value distributions that arise from the nature of the dimensionless shape parameter, κ – the Gumbel ($\kappa = 0$), the heavy-tailed Frechét ($\kappa > 0$), and the upper bounded Weibull distribution ($\kappa < 0$) distributions (Coles, 2001). The Cumulative Distribution Function (CDF) of the GEV distribution is given by:

$$F(x|\kappa, \beta, \alpha) = \exp\left(-\left(1 + \kappa\left(\frac{x-\alpha}{\beta}\right)\right)^{-1/\kappa}\right), 1 + \kappa\left(\frac{x-\alpha}{\beta}\right) \geq 0 \quad (1)$$

where α and β signify the location and scale parameter, respectively, in units of x . Here we use two different approaches to estimate extreme rainfall events: the stationary (ST) model, which considers the parameters of the GEV distribution constant over time; and the nonstationary (NS) model, in which the parameters of the extreme value distribution are considered as time-dependent (i.e., time is the covariate). We use a linear time-dependency between the location parameter (α) and time (t), i.e.,

$$\alpha_t = \alpha_0 + \alpha_1 t \quad (2)$$

where α_0 is the intercept and α_1 the slope of the linear model. Thus, the ST model has three unknown parameters to be inferred $\theta_s = \{\kappa, \beta, \alpha\}$; and the NS model has four unknown parameters, $\theta_n = \{\kappa, \beta, \alpha_0, \alpha_1\}$.

The quantile function of the GEV is given by (Coles, 2001):

$$x_p = \left(\left(-\frac{1}{\log(p)}\right)^{-\kappa} - 1\right) \left(\frac{\beta}{\kappa}\right) + \alpha_t, \kappa \neq 0 \quad (3)$$

where x_p is the quantile associated with the annual non-exceedance probability p . The extreme rainfall quantiles are estimated using the return period (Tr) as a proxy of non-exceedance probability, i.e., $Tr = 1/(1-p)$. Note that in the case of the NS model, the quantiles of the GEV will vary over time.

Bayesian inference is used for parameter and uncertainty estimation of ST and NS models. Within the framework of Bayesian inference for parameter estimation, Bayes' theorem becomes pivotal, allowing the update of the probability associated with a hypothesis – a model and its parameter values – by considering the available evidence, embodied as data. This approach provides a distribution of parameter values rather than point estimates, which is useful for hypothesis testing and uncertainty quantification (Luke et al., 2017). The posterior of parameter values of the model \mathcal{M}_j with parametrization θ_j , $P(\theta_j|\mathbf{X}, \mathcal{M}_j)$, is calculated from the prior distribution $P(\theta_j|\mathcal{M}_j)$, the likelihood function $L(\theta_j|\mathbf{X}, \mathcal{M}_j)$, and the evidence $P(\mathbf{X}|\mathcal{M}_j)$, i.e.,

$$P(\theta_j|\mathbf{X}, \mathcal{M}_j) = \frac{P(\theta_j|\mathcal{M}_j)L(\theta_j|\mathbf{X}, \mathcal{M}_j)}{P(\mathbf{X}|\mathcal{M}_j)} \quad (4)$$

where the subscript j denotes the model: $j = s$ for the ST model and $j = n$ for the NS model. \mathbf{X} is a vector of n observations $\mathbf{X} = \{x_1, \dots, x_n\}$ with the fitting period data of the AMF series (in our case, the full record for within-sample, and the first 30 observations for the out-of-sample). The prior distribution (Table 1) represents the knowledge of parameter values before data analysis. Here, we use a uniform distribution for the location (for both intercept and slope in the NS case) and scale parameters, in which the lower and upper limits of each parameter are set. For the shape parameter, there are some available priors in the literature, such as a Normal distribution with mean $\mu = 0.093$ and variance $\sigma^2 = 0.12^2$ (Papalexiou and Koutsoyiannis, 2013) based on regional information or a Beta distribution with mean $\mu = -0.10$ and variance $\sigma^2 = 0.122^2$ (Martins and Stedinger, 2000) based on knowledge of experts. These priors allow the sampling of negative values of the shape parameter. Negative values of the shape parameter imply an upper threshold in the GEV distribution, which is not realistic for rainfall, unless attributed to statistical variability (Deidda et al., 2021; Emma-nouil et al., 2020). Therefore, we use a uniform distribution in the range [0, 0.3], to avoid an upper threshold in the GEV distribution, as well as an infinite variance process (see also Hosking et al., 1985).

The likelihood function summarizes the information provided by the data. Mathematically, $L(\theta_j|\mathbf{X}, \mathcal{M}_j)$ is calculated using the probability density function of the GEV, i.e.,

$$L(\theta_j|\mathbf{X}, \mathcal{M}_j) = \prod_{i=1}^n P(X_i|\theta_j, \mathcal{M}_j) \quad (5)$$

The denominator of Bayes' theorem, also called the marginal probability or evidence, acts as a normalization constant to ensure that the posterior distribution has a unit area and is defined as:

$$P(\mathbf{X}|\mathcal{M}_j) = \int P(\theta_j|\mathcal{M}_j)L(\theta_j|\mathbf{X}, \mathcal{M}_j)d\theta_j \quad (6)$$

The parameter inference procedure can be made with the unnormalized values of the posterior distribution, i.e.,

$$P(\theta_j|\mathbf{X}, \mathcal{M}_j) \propto P(\theta_j|\mathcal{M}_j)L(\theta_j|\mathbf{X}, \mathcal{M}_j) \quad (7)$$

To approximate the posterior distribution of the parameter values, we use Markov chain Monte Carlo (MCMC) simulations with the Differential Evolution Adaptive Metropolis (DREAM_(ZS)) algorithm (Vrugt,

Table 1

Specification of the prior distribution of the GEV parameters of the stationary and nonstationary models. $\mathcal{U}(a,b)$ is the uniform distribution on the closed interval between a and b .

Parameter	Prior	a	b	Unit
κ	$\mathcal{U}(a,b)$	0	0.3	–
β	$\mathcal{U}(a,b)$	0	50	mm d ⁻¹
α_0	$\mathcal{U}(a,b)$	0	250	mm d ⁻¹
α_1	$\mathcal{U}(a,b)$	-5	5	mm d ⁻¹ y ⁻¹

2016). DREAM_(ZS) requires the setup of some parameters that depend on the case study: the dimension of the problem (number of parameters); the number of Markov chains; and the number of generations. We follow Luke et al. (2017) and generate 3 Markov chains with 8,000 samples each, which is enough to ensure the convergence of the chains in this case. Convergence metrics are monitored to determine when convergence of the sampled chains has been achieved to a limiting distribution. We use the last quarter of the sampled chains to summarize the posterior distribution of the GEV model parameters. This amounts to 6,000 posterior samples. We modify the source code available in the Support Information of Luke et al. (2017) to perform the parameter inference procedure described here for the GEV distribution.

2.3. Performance assessment

The posterior samples of the MCMC simulations with the DREAM_(ZS) algorithm are used for model comparison and selection. We evaluate the performance of the ST and NS models for within-sample prediction, i.e., ability of the model to reproduce the data used in the fitting period, and for out-of-sample prediction, in which data that are not used in the fitting procedure are used to assess model performance. To predict the out-of-sample data with the NS model, we follow the same methodology as used by Luke et al., 2017. The first approach extrapolates the trend of the fitting period to the evaluation period (NS model). In the second approach, the values of the NS model parameters at the end of the fitting period are used to predict the evaluation period. This so-called updated stationary or uST model, and assumes that the distribution of annual maximum precipitation is invariant in the evaluation period.

We use two different methods to assess the performance of the ST and NS models for the fitting (within-sample) and evaluation (out-of-sample) data sets. First, we assess the within-sample and out-of-sample model performance using the Continuous Ranked Probability Score (CRPS; Vrugt, 2023), i.e.,

$$\text{CRPS}_j = (\hat{\alpha} - x)(2F(x|\hat{\kappa}, \hat{\beta}, \hat{\alpha}) - 1) + \frac{\hat{\beta}}{\hat{\kappa}}(1 - (2 - 2^{\hat{\kappa}})\Gamma(1 - \hat{\kappa})) + \frac{2\hat{\beta}}{\kappa}(\Gamma(1 - \hat{\kappa}, -\log_e(F(x|\hat{\kappa}, \hat{\beta}, \hat{\alpha}))) - F(x|\hat{\kappa}, \hat{\beta}, \hat{\alpha})) \quad (8)$$

where $F(x|\hat{\kappa}, \hat{\beta}, \hat{\alpha})$ is the GEV CDF evaluated at $\hat{\theta}_j = \{\hat{\kappa}, \hat{\beta}, \hat{\alpha}\}$, which is the maximum a posteriori (MAP) parameter set, and Γ is the upper incomplete gamma function. The CRPS is a so-called scoring rule that summarizes the GEV distribution forecast (or simulation) of the MAP parameter values in a single reward-oriented value. Larger values of this scoring rule are preferred.

In the second approach, we assess the model's performance using the evidence (denominator) of Bayes' theorem (previously introduced in Equation (6)). This so-called marginal likelihood is equal to the probability that the model has generated the data for all possible values of the parameters and is often referred to as the model evidence. The evidence is a formalization of Occam's Razor in that a simple model is preferred unless a more complex model is significantly better at explaining the data. We compare different GEV models using the Bayes factor (BF):

$$B_{j,k} = 2\log_e\left(\frac{P(\mathbf{X}|\mathcal{M}_j)}{P(\mathbf{X}|\mathcal{M}_k)}\right) \quad (9)$$

where $B_{j,k}$ is the value of the BF of the model \mathcal{M}_j against the model \mathcal{M}_k . Log-scale formulation of the BF is chosen for a simpler interpretation: $\text{BF} > 0$ supports the selection of \mathcal{M}_j , while $\text{BF} < 0$ supports the selection of \mathcal{M}_k . According to Kass and Raftery (1995), values of $B_{j,k}$ between 0 and 2 indicate weak support, values between 2 and 6 a positive support, values between 6 and 10 a strong support, and values greater than 10 indicate a very strong support to the \mathcal{M}_j . The evidence of each model is estimated using the Gaussian Mixture Importance (GAME; Volpi et al., 2017) sampling, available in DREAM Package version 2.0. Essentially, GAME is

a Monte Carlo integration technique. The use of GAME requires the user to provide as input the posterior distribution – in our case for both fitting and evaluation periods. Therefore, for the within-sample frequency analysis, GAME estimates the model evidence by $P(\theta_j|\mathbf{X}, \mathcal{M}_j)$, and for the out-of-sample frequency analysis by $P(\theta_j|\mathbf{X}^*, \mathcal{M}_j)$ – in this case, it is necessary to specify the prior distribution of the evaluation period for the ST, uST and NS models. Similar to Luke et al. (2017), the prior distributions of the evaluation period of the ST, uST, and NS models are taken from the posterior distribution of the fitting period of each model by fitting a Gaussian Mixture Model (GMM) with m components. The GMM parameters are estimated by maximum likelihood, with the successive increase in the number of components until $\text{BIC}_m - \text{BIC}_{m+1} < 2$ using the Expectation Maximization (EM) algorithm, where BIC is the Bayesian information criterion (Schwarz, 1978).

2.4. Trend analysis

Here, we assume the presence of a significant trend as a proxy to the nonstationarity of the AMS. The trend analysis of the AMS is based on the posterior distribution of the slope of the location parameter of the GEV distribution (α_1). Trends are considered significant if the posterior distribution agrees with the signal of the change – in this case, if the zero value (no trend) is not included in the 95 % credibility interval of α_1 .

2.5. Magnification factors

The impact of the observed trends in rainfall AMS is estimated using the magnification factor (Vogel et al., 2011). The magnification factor (MF) is the ratio of extreme rainfall quantiles of some future year and some reference year, i.e.,

$$\text{MF} = \frac{x_p(t_0 + \Delta t)}{x_p(t_0)} \quad (10)$$

where x_p is the quantile function of the GEV distribution, t_0 is the reference year and Δt is some future planning horizon. As suggested by Vogel et al. (2011), we employ a decadal MF ($\Delta t = 10$ years). We estimate the MF for the 2-year quantile (median quantile, as a proxy of mean extreme rainfall behavior), and for the 100-year quantile, as a proxy for large rainfall events. The MF is estimated using the MAP parameter set inferred from the full record of each AMS.

3. Results and discussion

3.1. Performance evaluation

Here, we present the results of the models' performance assessment using the Continuous Ranked Probability Score (CRPS) evaluated with the MAP parameter set for both within-sample and out-of-sample prediction. Intuitively, we would expect a better performance of the NS model in nonstationary rainfall records (i.e., with significant trends). Thus, to test the hypothesis that the NS model would be preferable in nonstationary rainfall records, the rainfall records were initially separated in AMS with and without statistically significant trends in the fitting period (within-sample case), and with persistent trends (trend detected in both fitting period and full record for the out-of-sample case). Here, the presence of significant trends is detected by evaluating whether the zero value (i.e., no trend) is outside the 95 % credibility interval of the posterior distribution of the parameter α_1 (slope of the GEV location parameter). The results are summarized in Fig. 2. Considering the within-sample case (Fig. 2a), the NS model resulted in larger values of CRPS for most of the records when compared with the ST model – the NS model is preferred over the ST model in 78 % of the records. These results can be easily observed by the predominance of points above the 1:1 line. This is also the case when we compare the ST and NS models only for nonstationary records, in which the NS model is

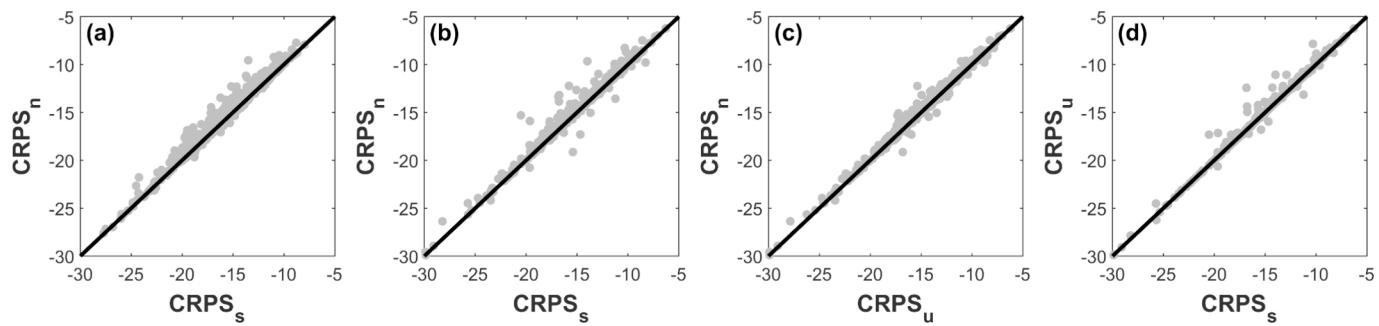


Fig. 2. Comparison of ST, uST and NS models using Continuous Ranked Probability Score (CRPS) evaluated with the MAP parameter set for (a) within-sample frequency analysis comparing ST and NS models, and (b-d) out-of-sample frequency analysis, comparing (b) ST and NS models, (c) uST and NS models, and (d) ST and uST models. The black line represents the 1:1 line.

supported in 94 % of the records – which is expected since the NS model is strongly preferred in most of the records. One of the main purposes of the NS models is to be used to estimate, for example, the distribution of extreme rainfall events in a future scenario. Therefore, it is important to test its out-of-sample prediction abilities, and not only within-sample prediction abilities. Fig. 2(b-d) summarizes the results for the out-of-sample prediction using the ST, uST, and NS models. Our results suggest that models with NS characteristics (i.e., uST and NS) show a better predictive ability over the ST model in roughly 86 % of the records (NS model in 65 % and uST in 21 %). Looking only at records with persistent trends, the models with nonstationary characteristics also outperform the ST model in 93 % of the records (NS model in 48 % and uST in 45 %). These results suggest that the NS model has a better performance compared to the ST model for both descriptive (reproduce historical data) and predictive (reproduce out-of-sample data) abilities and would be preferred to extreme rainfall frequency analysis.

The results derived from the within-sample and out-of-sample model performance assessment with the Bayes Factor (BF) using the full posterior samples derived from DREAM_(zs) are shown in Fig. 3. The BF is the ratio of two probabilities, which can take on a value between 0 and 1. By extension, the ratio of two probabilities cannot be negative. However, the log scale used here is symmetrical and allows for negative values ($BF < 0$, which is evidence against the alternative). For instance, a BF of -3 is as much evidence against it as a BF of $+3$ is evidence in favor of the hypothesis. Considering the within-sample case (Fig. 3a) and analyzing the record pool as a whole, can be seen that the support for the NS model decreased significantly when we used the BF for model selection, with the ST model showing better performance in 85 % of the records. Pooling the nonstationary records, the results suggest that the presence of significant trends in the fitting period tends to favor the selection of

the NS model in 58 % of the records (see markers with thick borders in Fig. 3a). For out-of-sample forecast Fig. 3 (b-d), the BF suggests that, overall, the ST model outperforms the uST and NS models in 56 % of the records. Even considering only records that exhibit persistent trends, extrapolation with the NS model is rarely preferred, with the NS model being selected only in 16 % of the records. In this case, the uST and ST models showed better forecast ability, with the uST model being selected in most of the records (57 %), followed by the ST model (27 %).

Our results suggest that there is a difference regarding the model selection when the CRPS and the BF are considered – it is a consequence of how these metrics work. It is important to note that the CRPS is a generalization of the absolute error (residual) in the context of descriptive or predictive distribution. Therefore, it accounts only for the ability of the model to reproduce the historical or out-of-sample data – the CRPS tends to support the selection of the NS model even when the AMS does not exhibit significant trends. In the case of no significant trends, predictions under ST and NS models are quite similar to each other. If models provide similar representations of the data, the simplest model should be chosen to avoid the selection of an overparameterized model. In addition, it is known that when we use more complex models, uncertainties arising from additional model complexity and epistemic uncertainty due to the use of the time-dependent model (i.e., the inability of the time to explain extreme rainfall variability) should be taken into account (Emmanouil et al., 2020; Luke et al., 2017; Serinaldi and Kilsby, 2015; Luke et al., 2017a). On the other hand, the BF can properly account for adding complexity and uncertainty, in which the simplest model is statistically preferred unless the more complex model is significantly better at explaining the data. Therefore, the use of CRPS tends to overestimate model complexity (i.e., tends to select the more complex model). According to our interpretation, the use of the BF for

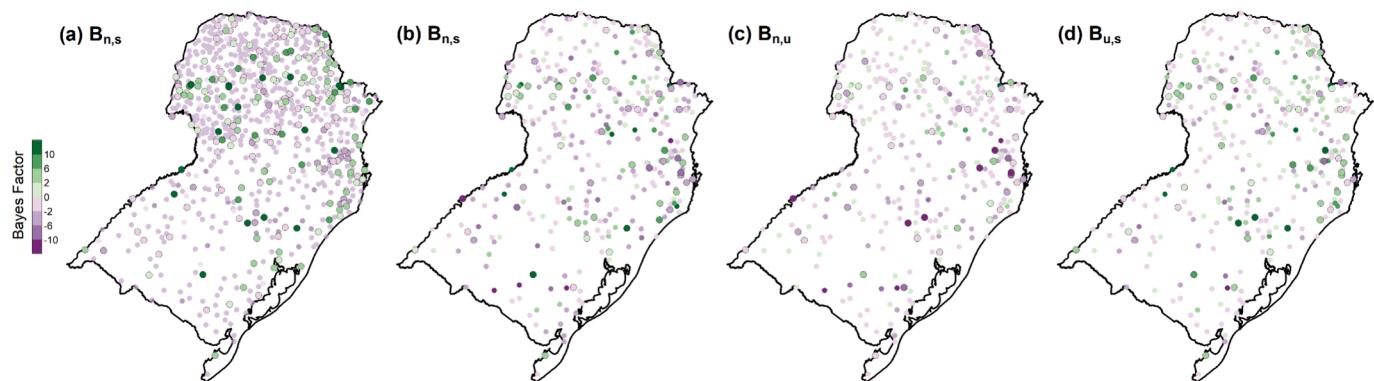


Fig. 3. Comparison of ST, uST and NS models using the Bayes Factor (BF) evaluated with the full posterior samples derived from DREAM_(zs) for (a) within-sample frequency analysis comparing ST and NS models, and (b-d) out-of-sample frequency analysis, comparing (b) ST and NS models, (c) uST and NS models, and (d) ST and uST models. Markers with thick borders indicate AMS with statistically significant trends in the fitting period (within-sample case), and with persistent trends (trend detected in both fitting period and full record for the out-of-sample case).

model selection provides a more consistent result – the BF demonstrates that the NS model rarely enhances sufficiently the statistical description of the out-of-sample data to justify the selection of a more complex model and associated prediction uncertainty (Luke et al., 2017). These findings lead to the main results of our work, in which the underlying uncertainties triggered by additional model complexity and epistemic uncertainties from the time-dependent structure prevent the selection of the NS model.

Although our results suggest that the uST model is the most appropriate model when there is a persistent nonstationary behavior in the AMS, it is not recommended its use to derive design rainfall or insurance analyses, as previously pointed out by Luke et al. (2017). Firstly, if the trends are indeed persistent over time, it is very unlikely that the predictions with the uST model to be accurate over a long period (in our case, about 30 years) – in this situation, despite the uncertainties and additional complexity issues, a safer prediction with the NS model can provide a more conservative design rainfall. Second, we only tested the ability of the models to predict the observed data of the fitting and evaluation periods, which does not ensure that the model predicts high return periods accurately – further efforts should be made to evaluate the accuracy of the models to predict specific return period events, in this case, focused on high quantiles with Monte Carlo simulations. (e.g., Martins and Stedinger, 2000, Yu et al., 2015). Therefore, due to the limitation of the NS model (i.e., its uncertainties) in predicting out-of-sample data and lack of knowledge about its accuracy in estimating high quantiles, we argue that its use should be restricted to detecting past changes or assessing current risks related to extreme precipitation with return period consistent with the length of the available data.

3.2. Extreme rainfall quantiles

Here, we present the results from the extreme rainfall quantiles under nonstationary assumption. Since the NS model structure can take into account the observed trend in the AMS, there are impacts on the estimated quantiles. These results are summarized using decadal magnification factors (i.e., the ratio of the 10-year future quantile relative to some reference year), estimated to 2- and 100-year extreme rainfall quantiles (Figs. 4–5). When we examine the record pool as a whole, there is a predominance of magnification factor values above 1 (i.e., future quantiles are higher than the reference period), with 73 % of the records exhibiting increases in extreme quantiles (blue tones in Fig. 4 and white histograms in Fig. 5). If we evaluate only AMS with significant trends, approximately 91 % of the records show increases in extreme

quantiles, which can be observed in the gray histograms in Fig. 5(a-b), with the magnification factor values shifted towards large values. Several studies also report that extreme rainfall in SB is increasing in magnitude, independently of the period analyzed and method utilized (e.g., Chagas and Chaffe, 2018; Doyle et al., 2011; Haylock et al., 2006; Liebmann et al., 2004; Naumann et al., 2012; Penalba and Robledo, 2010; Pinheiro et al., 2013; Re and Barros, 2009; da Teixeira, 2011). The sign of change for both return periods is the same, but the magnitude of the change is different, with the magnification factor values of the 2-year return period being slightly larger than the 100-year. This is a direct consequence of our model assumptions, in which the linear trend in the location parameter with fixed scale parameter results in a changing coefficient of variation – in this case, the ratio between quantiles in different return periods is not constant (Prosdocimi and Kjeldsen, 2021). Even though our model parameterization does not explicitly evaluate trends in different quantiles, the differences in the magnification factor can be expected since extreme events with different magnitudes could have different signs and rates of change due to different causing mechanisms. For example, Berg et al. (2013) and Fowler et al. (2021) outlined that there is evidence that long-duration extreme rainfall (1 day or more) increases with climate warming at approximately the Clausius–Clapeyron relationship (about 6–7 % K⁻¹). On the other hand, rare events can respond at higher rates, and short-duration extreme rainfall (hourly and sub-hourly) can respond more severely, about 2 times more (Fowler et al., 2021).

To further illustrate the results of the extreme rainfall quantiles, Fig. 6 portrays the return level estimates, reflecting the extreme rainfall quantiles evaluated at the most recent available year within the record (uST model), and for 15-year parameter extrapolation (NS model), for three specific rainfall gauges. These gauges serve as proxies in illustrating our findings regarding both ST and NS models. The results highlight important facets of the nonstationary models. Firstly, in some cases, the NS models exhibit notably wider credibility intervals, indicating higher uncertainty compared to the ST assumption. When a positive slope is observed in the GEV location parameter, the predictions derived from the NS model (Fig. 6, first to third frames) demonstrate larger return levels. Examining the distribution of return level values that delineates the probability density for a 50-year rainfall event (Fig. 6, fourth frame), a pattern emerges: under the ST model, extreme rainfall quantiles tend to concentrate around a common region – closer to the MAP estimates. In contrast, the NS model portrays a lower and more dispersed probability density, underscoring increased uncertainty and a tendency toward larger values. Conversely, in cases where

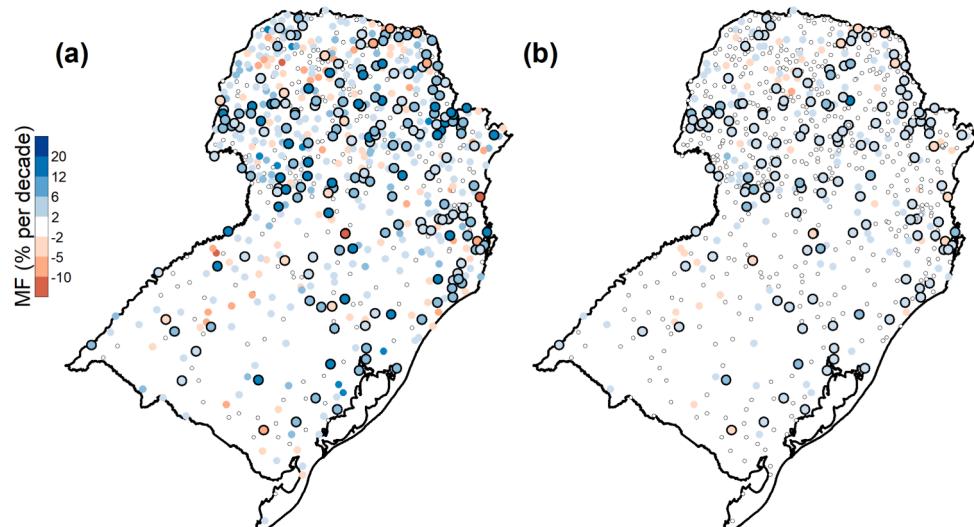


Fig. 4. Spatial distribution of decadal magnification factor values (in percentage) for (a) 2-year and (b) 100-year extreme rainfall quantiles. Large markers with thick borders show rainfall records with significant trends detected in the fitting period.

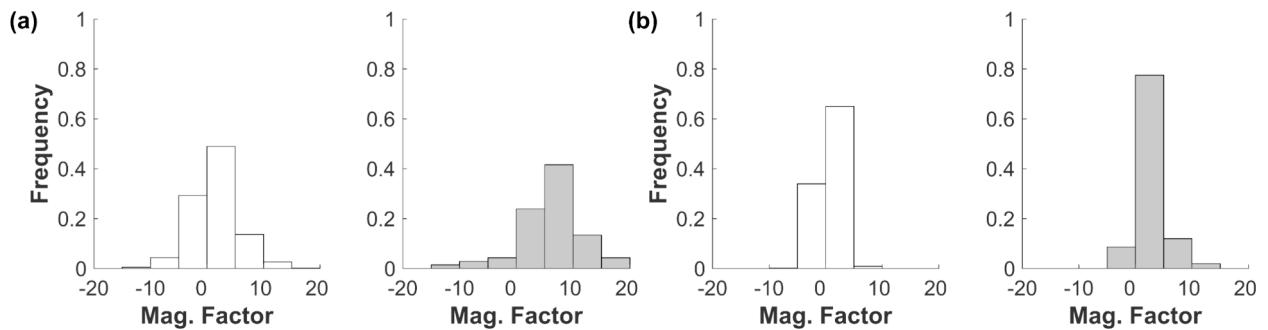


Fig. 5. Histograms of decadal magnification factor values for rainfall gauges without significant trend detected in the fitting period (white), and with significant trend detected in the fitting period (grey) for (a) 2-year and (b) 100-year extreme rainfall events.

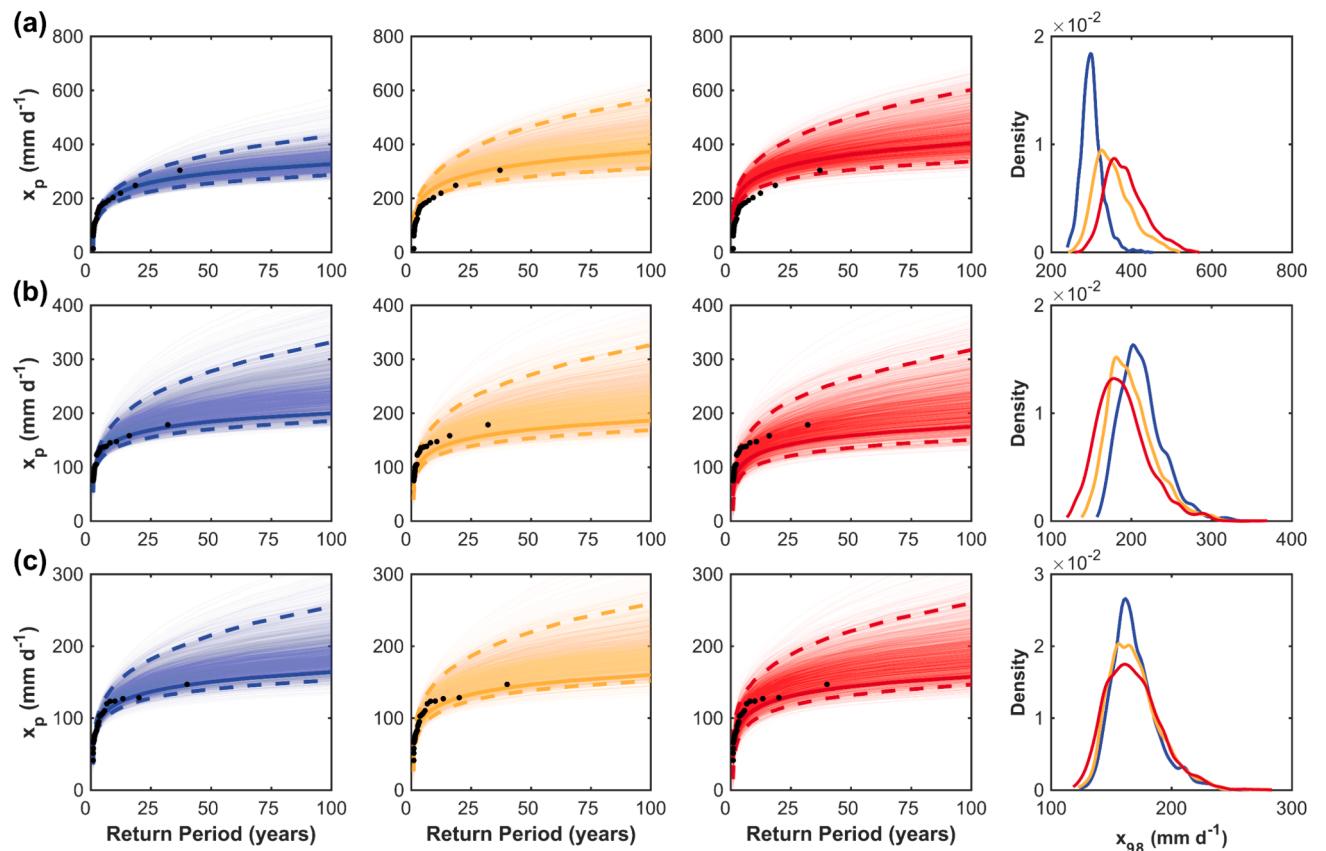


Fig. 6. (a-c) Frequency curves under ST (blue) and uST (yellow, extreme rainfall quantiles evaluated at last available year of the record), and NS (red, 15 years extrapolation) models, and probability density for a 50-year return period extreme rainfall event ($p = 98\%$) derived from the (a) Bananal rainfall record (ANA No. 02548043). (b) and (c) is the same as in (a), but for São Bernardo (ANA No. 02854013) and Triolância (ANA No. 02350052) rainfall records. The solid lines represent the maximum a posteriori (MAP) estimate of the posterior distribution, the dashed lines represent the 95 % credibility intervals, and the black circles represent the data under the empirical return period (stationary).

negative slopes are evident (Fig. 6b), careful attention is warranted. Beyond the heightened uncertainty associated with parameter inference, return level predictions under NS models are smaller, signifying a decreasing GEV location parameter over time. This assumption implies a higher risk, advocating caution in opting for the NS model in such scenarios.

For rainfall records devoid of significant trends (Fig. 6c), although estimates under ST and NS models might appear similar, nuanced differences emerge in the probabilistic distribution of quantiles. Here, the NS model demonstrates substantial uncertainty, as evidenced by a more dispersed density. This highlights the necessity for measures controlling complexity to steer clear of selecting a more intricate model with substantial uncertainty, especially when it yields a similar representation of

the data.

4. Impact of the observed trends in the parameter inference

Here, we used the simplest nonstationary model, i.e., with a linear trend in the location parameter. Although simple, this assumption is reasonable since a more complex trend model can easily result in an overfitted model. However, even considering the simplest trend model for the GEV location parameter, there are important implications in the inference of the scale parameter which is kept constant in the model, that was kept constant over time. Fig. 7 presents the values of the GEV distribution parameters for the MAP parameter set. The results suggest that the NS model tends to favor lower values of the scale parameter

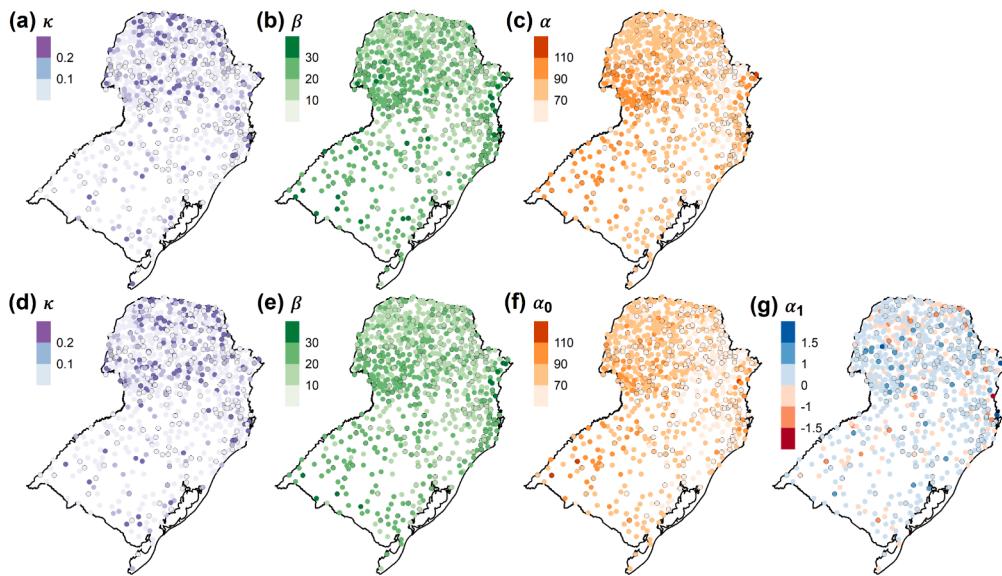


Fig. 7. Spatial distribution of MAP parameter set values estimated using the full record of each AMS for (a-c) stationary, and (d-g) nonstationary models. Large markers with thick borders show rainfall records with significant trends detected in the fitting period.

compared to the ST model, especially in the case of records with significant trends detected in the fitting period. This is a consequence of our model assumptions and the way that each model represents the data. For the ST model, the increase observed in the AMS (caused by the presence of a trend) is represented by a relatively large value of the scale parameter, which is simply a representation of the dispersion of the data. Under the NS model, the increase observed in the AMS is accounted for by an increase in the location parameter rather than the scale parameter. In the NS model, the scale parameter does not describe the dispersion of the sample as a whole, but rather the dispersion of the change in the location parameter over time. This result is also reported by Luke et al. (2017) and Prosdocimi and Kjeldsen (2021) for the Log-Pearson III and GEV distributions, respectively.

This concept is further illustrated in Fig. 8, which shows the posterior parameter distribution of a specific AMS that serves as a proxy of our findings regarding both ST and NS model's parameter inference. The posterior distribution of the scale parameter under the uST and NS model is shifted towards lower values compared with the NS model. Also, the posterior distribution of the location and shape parameters

under the nonstationary assumption show lower and more dispersed densities relative to the stationary assumption, especially for the location parameter – indicating higher uncertainties.

5. Conclusions

In this study, we use annual maximum rainfall records from Southern Brazil to test: (1) the suitability of NS models in the presence of statistically significant trends, and (2) their potential in modeling out-of-sample data to understand whether the presence of significant trends justifies the use of a nonstationary frequency analysis model. Employing a nonstationary model based on the generalized extreme value distribution, we allow the location parameter to dynamically change over time via a linear trend model, taking into account the underlying uncertainties in the model's performance and quantiles estimation.

Model selection based on maximum a posteriori parameter estimates may lead to the selection of a nonstationary model for within-sample and out-of-sample extreme rainfall prediction. However, taking into account the uncertainty associated with the estimated parameters, the

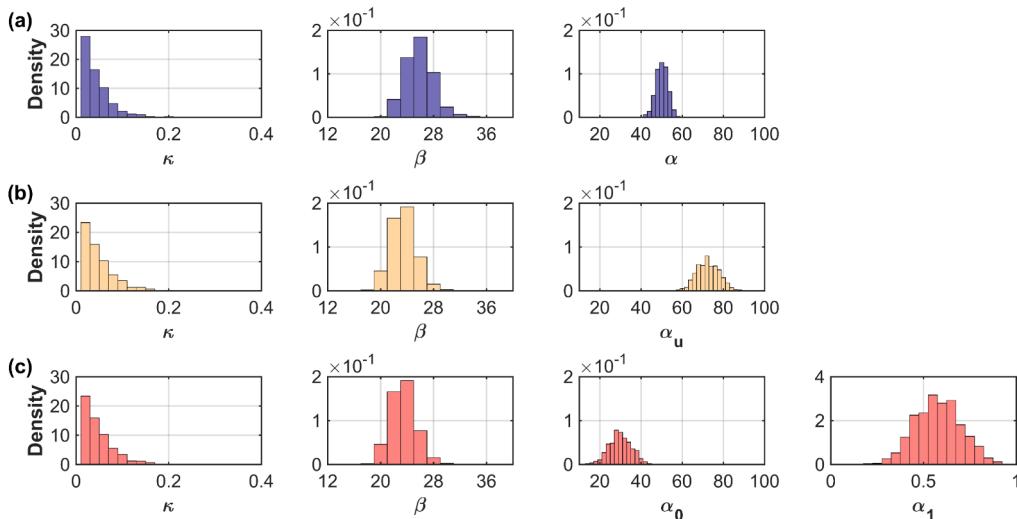


Fig. 8. Histogram of the posterior distribution of the estimated parameters under (a) ST (blue), (b) uST (yellow), and (c) NS (red) models for the Capela da Ribeira rainfall record (ANA No. 02449000).

NS model is rarely supported for prediction. The uncertainty triggered by the additional parameter and epistemic uncertainties due to the time-varying structure of the nonstationary model is the main limitation of the nonstationary assumption. We suggest that only the presence of significant trends is not sufficient to justify the selection of the NS model and that its use should be addressed with physical knowledge about the origin and persistence of the observed changes over time, properly taking into account the uncertainty resulting from additional parameters.

CRediT authorship contribution statement

Gabriel Anzolin: Formal analysis, Data curation, Conceptualization, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Debora Y. de Oliveira:** Conceptualization, Methodology, Writing – review & editing. **Jasper A. Vrugt:** Conceptualization, Methodology, Writing – review & editing.. **Amir AghaKouchak:** Conceptualization, Methodology, Writing – review & editing. **Pedro L.B. Chaffe:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- Anzolin, G., Chaffe, P.L.B., Vrugt, J.A., AghaKouchak, A., 2023. Using climate information as covariates to improve nonstationary flood frequency analysis in Brazil. *Hydrol. Sci. J.* 68, 645–654. <https://doi.org/10.1080/02626667.2023.2182212>.
- Bador, M., Donat, M.G., Geoffroy, O., Alexander, L.V., 2018. Assessing the Robustness of Future Extreme Precipitation Intensification in the CMIP5 Ensemble. *J. Clim.* 31, 6505–6525. <https://doi.org/10.1175/JCLI-D-17-0683.1>.
- Barbero, R., Fowler, H.J., Blenkinsop, S., Westra, S., Moron, V., Lewis, E., Chan, S., Lenderink, G., Kendon, E., Guerreiro, S., Li, X.-F., Villalobos, R., Ali, H., Mishra, V., 2019. A synthesis of hourly and daily precipitation extremes in different climatic regions. *Weather Clim. Extrem.* 26, 100219 <https://doi.org/10.1016/j.wace.2019.100219>.
- Barnett, T.P., Hasselmann, K., Chelliah, M., Delworth, T., Hegerl, G., Jones, P., Rasmusson, E., Roeckner, E., Ropelewski, C., Santer, B., Tett, S., 1999. Detection and attribution of recent climate change: A status report. *Bull. Am. Meteorol. Soc.* 80, 2631–2659. [https://doi.org/10.1175/1520-0477\(1999\)080<2631:DAAORC>2.0.CO;2](https://doi.org/10.1175/1520-0477(1999)080<2631:DAAORC>2.0.CO;2).
- Berg, P., Moseley, C., Haerter, J.O., 2013. Strong increase in convective precipitation in response to higher temperatures. *Nat. Geosci.* 6, 181–185. <https://doi.org/10.1038/ngeo1731>.
- Cavalcanti, I.F.A., Ferreira, N.J., da Dias, M.A.F.S., da Silva, M.G.A.J., 2009. *Tempo e clima no Brasil. Oficina de Textos, São Paulo*.
- Chagas, V.B.P., Chaffe, P.L.B., 2018. The role of land cover in the propagation of rainfall into streamflow trends. *Water Resour. Res.* 54, 5986–6004. <https://doi.org/10.1029/2018WR022947>.
- Chagas, V.B.P., L. B. Chaffe, P., Addor, N., M. Fan, F., S. Fleischmann, A., C. D. Paiva, R., Siqueira, V.A., 2020. CAMELS-BR: Hydrometeorological time series and landscape attributes for 897 catchments in Brazil. *Earth Syst Sci Data* 12, 2075–2096. <https://doi.org/10.5194/essd-12-2075-2020>.
- Cheng, L., AghaKouchak, A., 2014. Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Sci. Rep.* 4, 1–6. <https://doi.org/10.1038/srep07093>.
- Cheng, L., Phillips, T.J., AghaKouchak, A., 2015. Non-stationary return levels of CMIP5 multi-model temperature extremes. *Clim. Dyn.* 44, 2947–2963. <https://doi.org/10.1007/s00382-015-2625-y>.
- Coles, S., 2001. *An introduction to statistical modelling of extreme values*. Springer, London, UK.
- Cook, L.M., Anderson, C.J., Samaras, C., 2017. Framework for Incorporating Downscaled Climate Output into Existing Engineering Methods: Application to Precipitation Frequency Curves. *J. Infrastruct. Syst.* 23 [https://doi.org/10.1061/\(ASCE\)IS.1943-555X.0000382](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000382).
- Cook, L.M., McGinnis, S., Samaras, C., 2020. The effect of modeling choices on updating intensity-duration-frequency curves and stormwater infrastructure designs for climate change. *Clim. Change* 159, 289–308. <https://doi.org/10.1007/s10584-019-02649-6>.
- Teixeira, M. da S., Satyamurti, P., 2011. Trends in the Frequency of Intense Precipitation Events in Southern and Southeastern Brazil during 1960–2004. *J. Clim.* 24, 1913–1921. <https://doi.org/10.1175/2011JCLI3511.1>.
- Deidda, R., Hellies, M., Langousis, A., 2021. A critical analysis of the shortcomings in spatial frequency analysis of rainfall extremes based on homogeneous regions and a comparison with a hierarchical boundaryless approach. *Stoch. Env. Res. Risk a.* 35, 2605–2628. <https://doi.org/10.1007/s00477-021-02008-x>.
- Doyle, M.E., Saurral, R.I., Barros, V.R., 2011. Trends in the distributions of aggregated monthly precipitation over the La Plata Basin. *Int. J. Climatol.* 32, 2149–2162. <https://doi.org/10.1002/joc.2429>.
- Emmanouil, S., Langousis, A., Nikolopoulos, E.I., Agnastostou, E.N., 2020. Quantitative assessment of annual maxima, peaks-over-threshold and multifractal parametric approaches in estimating intensity-duration-frequency curves from short rainfall records. *J. Hydrol.* (amst.) 589, 125151. <https://doi.org/10.1016/j.jhydrol.2020.125151>.
- Emmanouil, S., Langousis, A., Nikolopoulos, E.I., Agnastostou, E.N., 2022. The Spatiotemporal Evolution of Rainfall Extremes in a Changing Climate: A CONUS-Wide Assessment Based on Multifractal Scaling Arguments. *Earths Future* 10. <https://doi.org/10.1029/2021EF002539>.
- Emmanouil, S., Langousis, A., Nikolopoulos, E.I., Agnastostou, E.N., 2023. Exploring the Future of Rainfall Extremes Over CONUS: The Effects of High Emission Climate Change Trajectories on the Intensity and Frequency of Rare Precipitation Events. *Earths Future* 11. <https://doi.org/10.1029/2022EF003039>.
- Fadhel, S., Rico-Ramirez, M.A., Han, D., 2017. Uncertainty of Intensity-Duration-Frequency (IDF) curves due to varied climate baseline periods. *J. Hydrol.* (amst.) 547, 600–612. <https://doi.org/10.1016/j.jhydrol.2017.02.013>.
- Fatici, S., Ivanov, V.Y., Paschalidis, A., Peleg, N., Molnar, P., Rimkus, S., Kim, J., Burlando, P., Caporali, E., 2016. Uncertainty partition challenges the predictability of vital details of climate change. *Earths Future* 4, 240–251. <https://doi.org/10.1002/2015EF000336>.
- Fischer, E.M., Beyerle, U., Knutti, R., 2013. Robust spatially aggregated projections of climate extremes. *Nat. Clim. Chang.* 3, 1033–1038. <https://doi.org/10.1038/nclimate2051>.
- Fowler, H.J., Ali, H., Allan, R.P., Ban, N., Barbero, R., Berg, P., Blenkinsop, S., Cabi, N.S., Chan, S., Dale, M., Dunn, R.J.H., Ekström, M., Evans, J.P., Fosser, G., Golding, B., Guerrero, S.B., Hegerl, G.C., Kahraman, A., Kendon, E.J., Lenderink, G., Lewis, E., Li, X., O’Gorman, P.A., Orr, H.G., Peat, K.L., Prein, A.F., Pritchard, D., Schär, C., Sharma, A., Stott, P.A., Villalobos-Herrera, R., Villarini, G., Wasko, C., Wehner, M.F., Westra, S., Whitford, A., 2021. Towards advancing scientific knowledge of climate change impacts on short-duration rainfall extremes. *Philosophical Transactions of the Royal Society a: Mathematical, Physical and Engineering Sciences* 379, 20190542. <https://doi.org/10.1098/rsta.2019.0542>.
- Gregersen, I.B., Sørup, H.J.D., Madsen, H., Rosbjerg, D., Mikkelsen, P.S., Arnbjerg-Nielsen, K., 2013. Assessing future climatic changes of rainfall extremes at small spatio-temporal scales. *Clim. Change* 118, 783–797. <https://doi.org/10.1007/s10584-012-0669-0>.
- Hailegeorgis, T.T., Alfredsen, K., 2017. Analyses of extreme precipitation and runoff events including uncertainties and reliability in design and management of urban water infrastructure. *J. Hydrol.* 544, 290–305. <https://doi.org/10.1016/j.jhydrol.2016.11.037>.
- Haughton, N., Abramowitz, G., Pitman, A., Phipps, S.J., 2015. Weighting climate model ensembles for mean and variance estimates. *Clim. Dyn.* 45, 3169–3181. <https://doi.org/10.1007/s00382-015-2531-3>.
- Haylock, M.R., Peterson, T.C., Alves, L.M., Ambrizzi, T., Anunciação, Y.M.T., Baez, J., Barros, V.R., Berlato, M.A., Bidegain, M., Coronel, G., Corradi, V., García, V.J., Grimm, A.M., Karoly, D., Marengo, J.A., Marino, M.B., Moncunill, D.F., Nechet, D., Quintana, J., Rebello, E., Rusticucci, M., Santos, J.L., Trebejo, I., Vincent, L.A., 2006. Trends in Total and Extreme South American Rainfall in 1960–2000 and Links with Sea Surface Temperature. *J. Clim.* 19, 1490–1512. <https://doi.org/10.1175/JCLI3695.1>.
- Hosking, J.R.M., Wallis, J.R., Wood, E.F., 1985. Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments. *Technometrics* 27, 251–261. <https://doi.org/10.1080/00401706.1985.10488049>.
- Huard, D., Mailhot, A., Duchesne, S., 2010. Bayesian estimation of intensity-duration-frequency curves and of the return period associated to a given rainfall event. *Stochastic Environmental Research and Risk Assessment* 24, 337–347. <https://doi.org/10.1007/s00447-009-0323-1>.
- Kendon, E.J., Ban, N., Roberts, N.M., Fowler, H.J., Roberts, M.J., Chan, S.C., Evans, J.P., Fosser, G., Wilkinson, J.M., 2017. Do Convective-Permitting Regional Climate Models Improve Projections of Future Precipitation Change? *Bull. Am. Meteorol. Soc.* 98, 79–93. <https://doi.org/10.1175/BAMS-D-15-0004.1>.
- Lee, O., Sim, I., Kim, S., 2020. Application of the non-stationary peak-over-threshold methods for deriving rainfall extremes from temperature projections. *J. Hydrol.* (amst.) 585, 124318. <https://doi.org/10.1016/j.jhydrol.2019.124318>.
- Liebmann, B., Vera, C.S., Carvalho, L.M.V., Camilloni, I.A., Hoerling, M.P., Allured, D., Barros, V.R., Baez, J., Bidegain, M., 2004. An Observed Trend in Central South American Precipitation. *J. Clim.* 17, 4357–4367. <https://doi.org/10.1175/3205.1>.
- Lopez-Cantu, T., Prein, A.F., Samaras, C., 2020. Uncertainties in future U.S. extreme precipitation from downscaled climate projections. *Geophys. Res. Lett.* 47 <https://doi.org/10.1029/2019GL086797>.
- Lucas, C., Timbal, B., Nguyen, H., 2014. The expanding tropics: A critical assessment of the observational and modeling studies. *Wiley Interdiscip. Rev. Clim. Change* 5 (1), 89–112. <https://doi.org/10.1002/wcc.251>.

- Luke, A., Vrugt, J.A., AghaKouchak, A., Matthew, R., Sanders, B.F., 2017. Predicting nonstationary flood frequencies: Evidence supports an updated stationarity thesis in the United States. *Water. Resour. Res.* 53, 5469–5494. <https://doi.org/10.1002/2016WR019676>.
- Martins, E.S., Stedinger, J.R., 2000. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water. Resour. Res.* 36, 737–744. <https://doi.org/10.1029/1999WR900330>.
- Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., 2008. Climate change: Stationarity is dead: Whither water management? *Science* 319 (5819), 573–574. <https://doi.org/10.1126/science.1151915>.
- Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., Dettinger, M.D., Krysanova, V., 2015. On Critiques of “stationarity is Dead: Whither Water Management?” *Water Resour. Res.* 51, 7785–7789. <https://doi.org/10.1002/2015WR017408>.
- Montanari, A., Koutsoyiannis, D., 2014. Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resour. Res.* 50, 9748–9756. <https://doi.org/10.1002/2014WR016092>.
- Moustakis, Y., Onof, C.J., Paschalis, A., 2020. Atmospheric convection, dynamics and topography shape the scaling pattern of hourly rainfall extremes with temperature globally. *Commun. Earth. Environ.* 1, 11. <https://doi.org/10.1038/s43247-020-0003-0>.
- Moustakis, Y., Papalexiou, S.M., Onof, C.J., Paschalis, A., 2021. Seasonality, intensity, and duration of rainfall extremes change in a warmer climate. *Earths Future* 9. <https://doi.org/10.1029/2020EF001824>.
- Naumann, G., Llano, P., Vargas, W.M., 2012. Climatology of the annual maximum daily precipitation in the La Plata Basin. *Int. J. Climatol.* 260, 247–260. <https://doi.org/10.1002/joc.2265>.
- Nerantzaki, S.D., Papalexiou, S.M., 2019. Tails of extremes: Advancing a graphical method and harnessing big data to assess precipitation extremes. *Adv. Water. Resour.* 134, 1–23. <https://doi.org/10.1016/j.advwatres.2019.103448>.
- Ouarda, T.B.M.J., Yousef, L.A., Charron, C., 2019. Non-stationary intensity-duration-frequency curves integrating information concerning teleconnections and climate change. *Int. J. Climatol.* 39, 2306–2323. <https://doi.org/10.1002/joc.5953>.
- Ouarda, T.B.M.J., Charron, C., St-Hilaire, A., 2020. Uncertainty of stationary and nonstationary models for rainfall frequency analysis. *Int. J. Climatol.* 40, 2373–2392. <https://doi.org/10.1002/joc.6339>.
- Pachauri, R.K., Allen, M.R., Barros, V.R., Broome, J., Cramer, W., Christ, R., 2014. Climate Change 2013 - The Physical Science Basis, Climate Change 2014: Synthesis Report. In: Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9781107415324>.
- Papalexiou, S.M., Koutsoyiannis, D., 2013. Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resour. Res.* 49, 187–201. <https://doi.org/10.1029/2012WR012557>.
- Papalexiou, S.M., Montanari, A., 2019. Global and regional increase of precipitation extremes under global warming. *Water Resour. Res.* 55, 4901–4914. <https://doi.org/10.1029/2018WR024067>.
- Papalexiou, S.M., AghaKouchak, A., Foufoula-Georgiou, E., 2018. A Diagnostic framework for understanding climatology of tails of hourly precipitation extremes in the United States. *Water Resour. Res.* 54, 6725–6738. <https://doi.org/10.1029/2018WR022732>.
- Penalba, O.C., Robledo, F.A., 2010. Spatial and temporal variability of the frequency of extreme daily rainfall regime in the La Plata Basin during the 20th century. *Clim. Change* 93, 531–550. <https://doi.org/10.1007/s10584-009-9744-6>.
- Pereima, M.F.R., Chaffé, P.L.B., de Amorim, P.B., Rodrigues, R.R., 2022. A systematic analysis of climate model precipitation in southern Brazil. *Int. J. Climatol.* 42 (8), 4240–4257. <https://doi.org/10.1002/joc.7460>.
- Petrow, T., Merz, B., 2009. Trends in flood magnitude, frequency and seasonality in Germany in the period 1951–2002. *J. Hydrol.* (amst.) 371, 129–141. <https://doi.org/10.1016/j.jhydrol.2009.03.024>.
- Pinheiro, A., Luiza, R., Graciano, G., Severo, D.L., 2013. TENDÊNCIA DAS SÉRIES TEMPORAIS DE PRECIPITAÇÃO DA REGIÃO SUL DO BRASIL Fundação Universidade Regional de Blumenau (FURB). Departamento De Engenharia Civil, Blumenau, SC, Brasil FURB, Departamento De Engenharia Ambiental, Blumenau, SC, Brasil FURB 281–290.
- Prein, A.F., Mearns, L.O., 2021. U.S. Extreme Precipitation Weather Types Increased in Frequency During the 20th Century. *J. Geophys. Res. Atmos.* 126. <https://doi.org/10.1029/2020JD034287>.
- Prein, A.F., Rasmussen, R.M., Ikeda, K., Liu, C., Clark, M.P., Holland, G.J., 2017. The future intensification of hourly precipitation extremes. *Nat. Clim. Chang.* 7, 48–52. <https://doi.org/10.1038/nclimate3168>.
- Prosdocimi, I., Kjeldsen, T., 2021. Parametrisation of change-permitting extreme value models and its impact on the description of change. *Stoch. Env. Res. Risk a.* 35, 307–324. <https://doi.org/10.1007/s00477-020-01940-8>.
- Ragno, E., AghaKouchak, A., Love, C.A., Cheng, L., Vahedifard, F., Lima, C.H.R., 2018. Quantifying changes in future intensity-duration-frequency curves using multimodel ensemble simulations. *Water Resour. Res.* 54, 1751–1764. <https://doi.org/10.1002/2017WR021975>.
- Ragno, E., AghaKouchak, A., Cheng, L., Sadegh, M., 2019. A generalized framework for process-informed nonstationary extreme value analysis. *Adv. Water. Resour.* 130, 270–282. <https://doi.org/10.1016/j.advwatres.2019.06.007>.
- Re, M., Barros, V.R., 2009. Extreme rainfalls in SE South America. *Clim. Change* 96, 119–136. <https://doi.org/10.1007/s10584-009-9619-x>.
- Read, L.K., Vogel, R.M., 2015. Reliability, return periods, and risk under nonstationarity. *Water Resour. Res.* 51, 6381–6398. <https://doi.org/10.1002/2015WR017089>.
- Roderick, T.P., Wasko, C., Sharma, A., 2020. An improved covariate for projecting future rainfall extremes? *Water Resour. Res.* 56, 2–11. <https://doi.org/10.1029/2019WR026924>.
- Schwarz, G., 1978. Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. <https://doi.org/10.1214/aos/1176344136>.
- Serago, J.M., Vogel, R.M., 2018. Parsimonious Nonstationary Flood Frequency Analysis. *Adv. Water Resour.* <https://doi.org/10.1016/j.advwatres.2017.11.026>.
- Serinaldi, F., Kilsby, C.G., 2015. Stationarity is undead: Uncertainty dominates the distribution of extremes. *Adv. Water Resour.* 77, 17–36. <https://doi.org/10.1016/j.advwatres.2014.12.013>.
- Šraj, M., Viglione, A., Parajka, J., Blöschl, G., 2016. The influence of non-stationarity in extreme hydrological events on flood frequency estimation. *Journal of Hydrology and Hydromechanics* 64, 426–437. <https://doi.org/10.1515/johh-2016-0032>.
- Touma, D., Stevenson, S., Swain, D.L., Singh, D., Kalashnikov, D.A., Huang, X., 2022. Climate change increases risk of extreme rainfall following wildfire in the western United States. *Sci. Adv.* 8. <https://doi.org/10.1126/sciadv.abm0320>.
- Veneziano, D., Langousis, A., Furcolo, P., 2006. Multifractality and rainfall extremes: A review. *Water Resour. Res.* 42 (June), 1–18. <https://doi.org/10.1029/2005WR004716>.
- Vogel, R.M., Yaindl, C., Walter, M., 2011. Nonstationarity: Flood Magnification and Recurrence Reduction Factors in the United States. *JAWRA Journal of the American Water Resources Association* 47, 464–474. <https://doi.org/10.1111/j.1752-1688.2011.00541.x>.
- Volpi, E., Schoups, G., Firmani, G., Vrugt, J.A., 2017. Sworn testimony of the model evidence: Gaussian Mixture Importance (GAME) sampling. *Water Resour. Res.* 53 (7), 6133–6158. <https://doi.org/10.1002/2016WR020167>.
- Vrugt, J.A., 2016. Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environ. Model. Softw.* 75, 273–316. <https://doi.org/10.1016/j.envsoft.2015.08.013>.
- Vrugt, Jasper A., 2023. The Promise of Diversity: Distribution-based Hydrologic Model Evaluation and Diagnostics. *ESS Open Archive.* 10.22541/essoar.168056799.99219343/v1.
- Vu, T.M., Mishra, A.K., 2019. Nonstationary frequency analysis of the recent extreme precipitation events in the United States. *J. Hydrol.* (amst.) 575, 999–1010. <https://doi.org/10.1016/j.jhydrol.2019.05.090>.
- Yu, X., T. A. Cohn, J. R. Stedinger, K. Karvazy, and V. Webster, 2015. Flood frequency analysis in the context of climate change. *World Environmental and Water Resources Congress*, pp. 2376–2385, Am. Soc. of Civ. Eng., Reston, Va.