

Research

Assessment of alternative methods for analysing maximum rainfall spatial data based on generalized extreme value distribution

Thales Rangel Ferreira¹  · Gilberto Rodrigues Liska²  · Luiz Alberto Beijo³ 

Received: 22 July 2023 / Accepted: 9 January 2024

Published online: 29 January 2024

© The Author(s) 2024 [OPEN](#)

Abstract

The present study aimed to analyze and spatially model maximum rainfall in the southern and southwestern regions of Minas Gerais using spatial statistical methods. Daily data on maximum rainfall were collected from 29 cities in the region. To obtain predictions of maximum rainfall for return periods of 2, 5, 10, 50, and 100 years, Bayesian Inference was employed, utilizing the most appropriate prior for each locality. The spatial analysis of the phenomenon based on results obtained through Bayesian Inference was conducted using interpolation methods, including Inverse Distance Weighting (IDW) and Kriging (Ordinary Kriging (OK) and Log-Normal Kriging (LK)). Different semivariogram models were used, and the most suitable one was selected based on cross-validation results for each method, which were also compared to those of IDW. Additionally, a spatial analysis was carried out using max-stable processes and spatial Generalized Extreme Value (GEV) distribution, with the models evaluated based on Takeuchi's Information Criteria. All models were also assessed by calculating the mean prediction error for six locations that were not used in model fitting. The results indicated that the most suitable models among Kriging and IDW for return periods of 2, 5, and 10 years were Gaussian (LK), Spherical (OK), and Wave (OK), respectively. Among the max-stable models and spatial GEV, the most suitable for modeling was the Smith max-stable model. Consequently, for spatial prediction over 50- and 100-year return periods, OK (Wave) and the Smith max-stable model were employed.

Article Highlights

- The non-informative prior distribution showed better results for more cities compared to the informative one.
- Techniques from the Theory of Extreme Values, Bayesian Inference and Geostatistics were applied jointly.
- Predictions of up to approximately 180 mm of daily rainfall were obtained.

Keywords GEV distribution · Kriging · Max-stable processes · Priors · Semivariogram

✉ Gilberto Rodrigues Liska, gilbertoliska@ufscar.br; Thales Rangel Ferreira, thales.rangel8@gmail.com; Luiz Alberto Beijo, prof.beijo@gmail.com | ¹Federal University of Alfenas, Rua Gabriel Monteiro da Silva, 700, Alfenas, MG CEP 37130-001, Brazil. ²Department of Agroindustrial Technology and Rural Socioeconomics, Federal University of São Carlos, Araras, Rod. Anhanguera km 174-SP-330, Araras, SP CEP 13600-970, Brazil. ³Department of Statistics, Federal University of Alfenas, Rua Gabriel Monteiro da Silva, 700, Alfenas, MG CEP 37130-001, Brazil.



1 Introduction

The occurrence of extreme climatic events has emerged as a mounting concern within both society and the scientific community. The frequency and repercussions of such events, including economic and societal damages, have garnered attention due to the adversities stemming from natural disasters [1]. These disasters are typically triggered by severe climatic factors, with intense rainfall being recognized as one of the primary causes.

The impacts caused by extreme weather events affect not only the occurrence of natural disasters, but also agricultural and livestock activities. During periods of heavy rains, the planting of crops with more fragile structures may be affected due to the decrease in soil oxygen quantity, which affects nutrient absorption [2]. In addition, heavy rainfall can result in adverse effects on public health, since exposure to flood and flood-prone areas can increase the risk of infectious diseases, such as leptospirosis, transmitted by contact with water or mud contaminated by animal urine, mainly rodents [3].

The study of the behavior of extreme variables, such as maximum rainfall, is conducted through approaches from the Extreme Values Theory. This theory was propelled by the work of [4], which described the three types of asymptotic extreme values distributions (Types I, II, and III). In this theory, the analysis of datasets with small sample sizes is common due to the nature of extreme variables. In this scenario, using Bayesian Inference for fitting extreme value models emerges as an effective approach for analyses, as it enables the incorporation of prior information, which can reduce uncertainties surrounding the parameters under study [5].

Based on point predictions obtained through Bayesian inference for different locations and return periods in a given region, it becomes possible to analyze the variable from a spatial perspective. In this study, various spatial statistical techniques were applied to investigate the spatial behavior of extreme rainfall in the southern and southwestern regions of Minas Gerais. These techniques focus on modeling variables that exhibit spatial structure, enabling inferences and predictions of attributes for unsampled locations within the spatial domain.

Spatial statistics use interpolation methods, which rely on known point values to estimate values at unsampled points [6]. Different spatial interpolation methods can be classified into two categories: Deterministic and Stochastic (e.g., Geostatistical). The Geostatistical approach is particularly appealing as it allows for the incorporation of complementary information from spatial covariates using multivariate methods [7]. In this study, both Deterministic and Geostatistical interpolation methods were employed, specifically the Inverse Distance Weighting (IDW), the Ordinary Kriging (OK) and Log-Normal Kriging (LK) methods.

Due to the extreme nature of the phenomenon under study, in addition to the Kriging and IDW methods, methodologies specific to the analysis of spatial extreme variables were employed, such as max-stable processes and the spatial Generalized Extreme Value (GEV) distribution. Through max-stable processes and the spatial GEV model, it is possible to integrate geographic and climatic covariates into trend surfaces, allowing for the spatial variation of extreme model parameters in the analyzed region and the spatial prediction of the attribute [8].

Regarding the spatial analysis of rainfall in the region of Minas Gerais, few studies in the literature have applied Geostatistical methods for interpolating the variable, only the studies by [9, 10] were found. However, distinguishing itself from the present study, these works focused on assessing the spatial distribution of the observed variable or the mean rainfall, without addressing predictions for specific return periods. It is worth noting that this region is considered the second largest mesoregion in the state of Minas Gerais, with 2.7 million inhabitants [11], an approximate size of the population of a country, such as, for example, Uruguay, which, according to the last census of 2011 [12], it is 3.39 million inhabitants. Furthermore, the majority of the state's coffee production is located in this region, with production of 13,684 thousand bags, which corresponds to approximately 51% of the State's total production [11].

Due to the possible consequences of intense rainfall and its influence on different socioeconomic and environmental aspects, it becomes necessary to study and understand this phenomenon through a spatial approach. Therefore, this study aimed to spatially model the maximum rainfall in the southern and southwestern region of Minas Gerais using different spatial methodologies applied from point rainfall predictions for different return periods obtained through both Bayesian inference and max-stable methods.

2 Materials and methods

2.1 Dataset description and region of study

The southern and southwestern region of Minas Gerais is one of the 12 mesoregions of the state and has an area of 49,524 km². According to [13], the climates present in the region, according to the Köppen classification, are Aw, Cwa, and Cwb. The first one is characterized by a winter with low rainfall and a rainy summer. Cwa and Cwb climates have milder temperatures compared to the Aw climate and are influenced by the high altitudes of the southern region of Minas Gerais. The Aw climate is mainly found in the northern portion of the region, while the Cwa climate is present in areas with higher altitudes, such as the Serra da Mantiqueira and nearby regions, and the Cwb climate is observed in the southern portion of the region, near the border with São Paulo, according to [14].

In this study, daily rainfall data, expressed in water depth (mm), were used for 29 cities in the southern and southwestern region of Minas Gerais, as highlighted in Fig. 1.

Daily records were obtained through the National Water Agency (ANA) and National Institute of Meteorology (INMET). The rainfall data refers to the years from 1980 to 2021. The daily rainfall data for each city were grouped into annual series, and from these series the corresponding maximum values for each year were selected. Thus, 29 maximum value series were obtained, each containing 42 data points, representing the maximum annual daily rainfall values for each city under study.

In order to evaluate the models used in this study, the maximum value series were separated into two distinct blocks. The first block, consisting of data from 1980 to 2011, was used for fitting the models. That is, the GEV distribution was fitted through Bayesian Inference and rainfall predictions were obtained for return periods of 2, 5, and 10 years for each city. Spatial modeling was then performed through IDW and Kriging from these predictions, obtaining spatial predictions.

The second block, with records from the period of 2012 to 2021, was used to evaluate the effectiveness of each model. For this, cross-validation methods and calculation of the mean prediction error (MPE) were used for locations not used in the model fitting.

Table 1 presents data from the meteorological stations used in the study. It is important to note that some of the stations had gaps in their records, which led to the use of data from other stations located in the same city, in order to

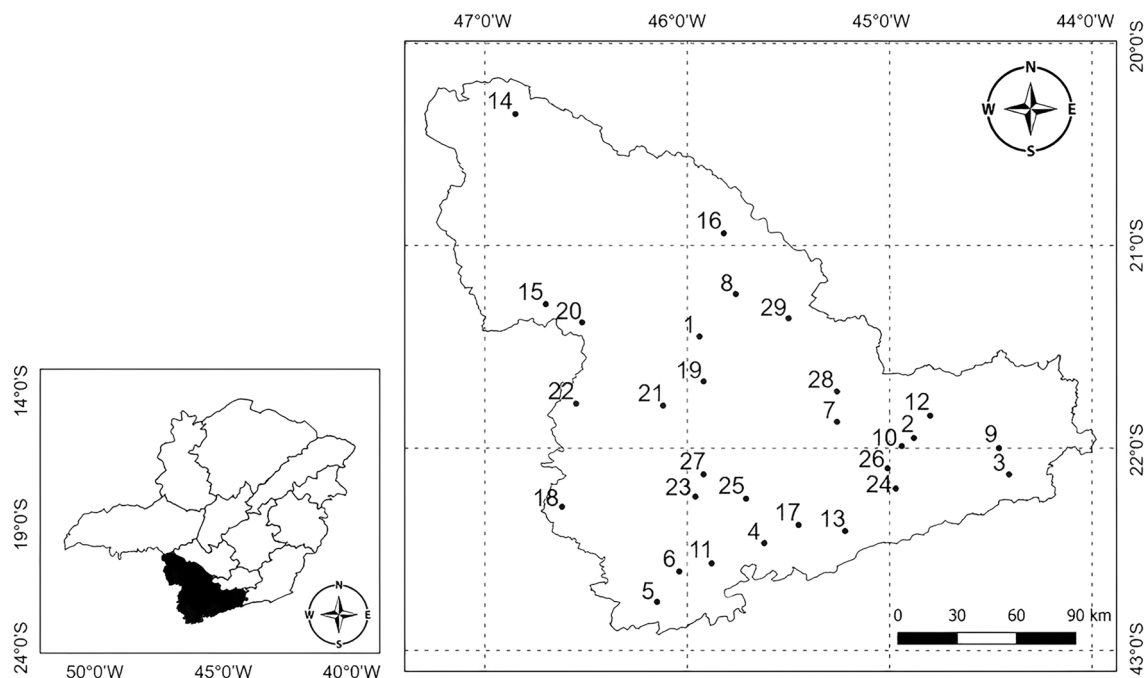


Fig. 1 Location of the meteorological stations of the 29 cities in southern and southwestern Minas Gerais analyzed in the study numbered in alphabetical order

Table 1 Data from meteorological stations of the 29 cities in southern and southwestern Minas Gerais analyzed, with their respective percentage of failures (PF)

Number	Order	Cities	Lat	Lon	Alt	PF(%)
2145042	1	Alfenas	-21.45	-45.94	888	7.14
2144004	2	Baependi	-21.95	-44.88	906	0.00
2244057	3	Bocaina de Minas	-22.13	-44.41	1180	0.00
2245070	4	Brazópolis	-22.47	-45.62	870	0.00
2246057	5	Camanducaia	-22.76	-46.15	1006	0.00
2246050	6	CambuÍ	-22.61	-46.04	868	7.14
2145008	7	Cambuquira	-21.87	-45.26	863	0.00
2145041	8	Campos Gerais	-21.24	-45.76	883	7.14
2144025	9	Carvalhos	-22.00	-44.46	1306	4.76
2144003	10	Caxambu	-21.99	-44.94	958	2.38
2245084	11	Consolação	-22.57	-45.88	979	0.00
2144037	12	Cruzília	-21.84	-44.80	1088	4.76
2245087	13	Delfim Moreira	-22.41	-45.22	1405	0.00
2046009	14	Delfinópolis	-20.35	-46.85	687	35.71
2146026	15	Guaxupé	-21.29	-46.70	837	0.00
2045026	16	Illicínea	-20.94	-45.82	900	7.14
2245083	17	Itajuba	-22.38	-45.45	851	0.00
2246052	18	Jacutinga	-22.29	-46.62	865	0.00
2145002*	19	Machado	-21.67	-45.92	818	11.9
2146030	20	Muzambinho	-21.38	-46.52	1064	0.00
2146029	21	Poço Fundo	-21.79	-46.12	944	0.00
2146048*	22	Poços de Caldas	-21.78	-46.55	1339	14.29
2245077	23	Pouso Alegre	-22.24	-45.96	891	14.29
2244071*	24	Pouso Alto	-22.20	-44.97	940	2.38
2245000	25	Santa Rita do SapucaÍ	-22.25	-45.71	824	7.14
2245081*	26	São Lourenço	-22.10	-45.01	1014	7.14
2245085*	27	Silvianópolis	-22.13	-45.92	851	4.76
2145003	28	Três Corações	-21.72	-45.26	847	0.00
2145043	29	Três Pontas	-21.36	-45.50	903	9.52

Cities where records from more than one weather station were used (*); Latitude (Lat); Longitude (Lon); Altitude (Alt)

complete the series. The cities where gaps were identified in the stations and the alternative stations used to supplement the series are: Machado (2145002-ANA, 2145033-ANA, and 83683-INMET), Poços de Caldas (2146048-ANA, 2146074-ANA, and 83681-INMET), Pouso Alto (2244071-ANA and 2244063-ANA), São Lourenço (83736-INMET, 2245081-ANA, and 2245107-ANA), Silvianópolis (2245085-ANA and 2245075-ANA).

Due to the failures in the records of some datasets obtained from INMET and ANA, even with the combination of data from different stations, some gaps persisted. To address this issue, the *k*-Nearest Neighbors imputation method was applied, which involves filling in the missing values from the weighted average of the *k*-nearest neighbors of that missing value [15]. In this way, complete datasets were obtained for all the studied locations.

In order to perform a descriptive analysis of the maximum rainfall data for each of the 29 study locations, the position measures of the series (mean, median, maximum and the coefficient of variation) were estimated.

2.2 Bayesian approach to maximum rainfall analysis

To verify the assumption of independence of the constructed maximum series and the presence of trend, the Ljung-Box test [16] and Mann-Kendall test [17, 18] were used, respectively, at a 5% significance level. Kolmogorov-Smirnov test was used for evaluate the goodness of fitting GEV distributions at a 5% significance level.

For each of the 29 maximum rainfall series from cities in the south and southwest of Minas Gerais, GEV distribution adjustments were performed in order to model the distribution of annual maximum rainfall. The GEV distribution

has parameters μ , σ , and ξ , which are, respectively, the position, scale, and shape parameters, and its probability density function is given by:

$$f(x) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\left(\frac{1+\xi}{\xi}\right)} \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}, \quad (1)$$

where, $1 + \xi \left(\frac{x - \mu}{\sigma} \right) > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$, and $\xi \in \mathbb{R}^*$. The form of the GEV distribution varies according to the value of the parameter ξ . For $\xi < 0$, the function is defined for $-\infty < x < \mu - \frac{\sigma}{\xi}$. For $\xi > 0$, the function is defined for $\mu - \frac{\sigma}{\xi} < x < \infty$. When ξ equals zero, the function is defined for $x \in \mathbb{R}$.

Several methods can be used to estimate the parameters of the GEV distribution, including Bayesian Inference. The Bayesian approach incorporates prior information about the phenomenon of interest and the data information to obtain the posterior density of the parameters. The prior information is represented by a prior distribution of the parameters, which reflects prior knowledge about them before obtaining or collecting the data. The information from the data is represented by the likelihood function. Thus, the posterior distribution contains all available parameter information, rendering it information-rich [19].

By using the prior distribution, $p(\theta)$, and the likelihood function, $L(\theta|\mathbf{x})$, it is possible to obtain the posterior distribution of θ , $p(\theta|\mathbf{x})$:

$$p(\theta|\mathbf{x}) \propto L(\theta|\mathbf{x})p(\theta), \quad (2)$$

where, \propto represents proportionality and $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ is the data sample. The likelihood function of the GEV distribution is given by:

$$L(\theta|\mathbf{x}) = \frac{1}{\sigma^n} \prod_{i=1}^n \left\{ \left[1 + \xi \left(\frac{x_i - \mu}{\sigma} \right) \right]^{-\left(\frac{1+\xi}{\xi}\right)} \right\} \exp \left\{ \sum_{i=1}^n \left\{ - \left[1 + \xi \left(\frac{x_i - \mu}{\sigma} \right) \right]^{-\left(\frac{1}{\xi}\right)} \right\} \right\}, \quad (3)$$

where, $\theta = (\mu, \sigma, \xi)$ and n is the sample size.

The prior distribution can be categorized into two distinct types based on the degree of prior knowledge available about the parameters θ . These types include the informative prior distribution and the non-informative prior distribution. When experts have prior information and prefer a specific range of one or more parameter values in a parameter space, the prior distribution is considered informative. On the other hand, when the researcher does not prefer any parameter value in a parameter space, the prior distribution is defined as non-informative.

[20] formulated a prior distribution for the three parameters of the GEV distribution with the aim of improving the estimation of these parameters, proposing that they follow a Trivariate Normal distribution, given by:

$$p(\theta) = (2\pi)^{1/2} |\Sigma_0|^{1/2} \exp \left\{ -\frac{1}{2} (\theta - \Phi_0)^t \Sigma_0^{-1} (\theta - \Phi_0) \right\}, \quad (4)$$

where, $\theta = (\mu; \log(\sigma); \xi)$ is the parameter vector, Φ_0 is the hyperparameter means vector, and Σ_0 is the hyperparameter variance-covariance matrix.

The information for the hyperparameters of the informative prior distribution for each analyzed location, that is, for the informative priors of each city, was obtained by fitting the GEV distribution to the data from the nearest municipality among the 28 available. For the non-informative prior, the respective values of the hyperparameters were adopted: $\Phi_0 = (0; 0; 0)$ and $\Sigma_0 = (10, 000; 10, 000; 100)$.

According to [21], the marginal distributions of the GEV parameters cannot be obtained analytically, necessitating the use of simulation techniques based on probability distributions. One such technique is Markov Chain Monte Carlo (MCMC) simulation, which utilizes iterative algorithms such as Metropolis-Hastings and Gibbs Sampler to generate variables from unknown density functions. This method aims to construct a density for each parameter, however, it is crucial to verify the convergence of the generated chains.

These algorithms can obtain samples from the posterior distribution by running a constructed Markov Chain, and the algorithm simulates Markov chains within a Monte Carlo integration to generate a set of points whose distribution converges to the *posterior* distribution [22].

The joint posterior distributions for the GEV distribution were obtained numerically using the Markov Chain Monte Carlo method. In this method, for each prior structure used, simulations with 120,000 iterations were performed, burning the first 20,000 iterations and adopting a thinning of every 20 iterations.

Several criteria have been proposed to assess the convergence of the posterior chains of the parameters. Among these methods, [23] recommends the application of the Raftery-Lewis [24], Geweke [25], and Heidelberg-Welch criteria [26]. Therefore, in this study, these criteria were employed to evaluate the convergence of the obtained chains.

In order to assess the performance of the different *prior* distributions, the return levels were analyzed for return periods of 2, 5, and 10 years. In this study, the return levels correspond to the predictions of maximum rainfall and are calculated as follows:

$$\hat{\chi}(T) = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left\{ 1 - \left[-\ln \left(1 - \frac{1}{T} \right) \right]^{-\hat{\xi}} \right\}. \quad (5)$$

where, $\hat{\chi}(T)$ represents the return level for a return period T , and μ , σ , and ξ are the posterior means of each parameter, obtained through Bayesian Inference

To evaluate the performance of the models using different priors, the results of the mean prediction error were used, given by:

$$MPE = \frac{1}{m} \sum_i^m \left| \frac{o_i - \hat{o}_i}{o_i} \right| \times 100, \quad (6)$$

where, \hat{o}_i is the estimated maximum rainfall for the i -th return period, o_i is the observed maximum rainfall from the historical series, and m is the number of predictions.

The predictions of rainfall for the 29 locations were obtained for the return periods of 2, 5, 10, and also for 50 and 100 years, and for these last two, the most suitable interpolation method among those tested was applied to obtain the prediction maps.

2.3 Spatial analysis from bayesian predictions

Through obtaining the predicted rainfall values and their corresponding coordinates for each meteorological station, the analysis can be extended from a spatial perspective. For this study, the classical IDW method and the Geostatistical methods were employed. For an unobserved location, the prediction of the attribute using IDW is given by:

$$\hat{z}(u_j) = \frac{\sum_{i=1}^n \frac{1}{d_{ij}^p} z(u_i)}{\sum_{i=1}^n \frac{1}{d_{ij}^p}}, \quad (7)$$

where, $\hat{z}(u_j)$ is the estimate of the attribute for a non-sampled location u_j , $z(u_i)$ is the observed value of maximum rainfall for an observed point, d_{ij} is the Euclidean distance between the i -th neighbor and point u_j , p is the power parameter, and n is the number of sampled points used for estimation.

The interpolation of values is performed through a weighted average of the observed values within a certain distance, where the closest observed values have greater weight ($1/d_{ij}^p$) in the weighted average. In this work, a value of $p = 2$ was used as the power parameter for the calculation of the weighted average. The IDW interpolator is widely employed in the analysis of geospatial data due to its simplicity and reliance solely on observed values and distances between points. Given its ease of understanding and application, IDW is often used for spatialization of variables at unsampled locations, making it one of the most commonly utilized methods alongside Kriging [27].

In contrast to IDW, Kriging methods are based on Geostatistical models that take into account the spatial correlation structure of the data. It is based on the assumptions of unbiasedness in estimation and minimum variance of estimates, and employs semivariogram functions to detect the spatial variability structure of the variable of interest [28–30].

Regarding semivariogram estimators, the classical estimator of Matheron is the most commonly used [31]. However, this estimator is sensitive to the presence of outliers and asymmetry of the random variable. Therefore, as an alternative estimator for problems where the data deviate from normality, the robust estimator of Cressie-Hawkins is presented, since, according to [32], it is not sensitive to the presence of outliers. The Cressie-Hawkins robust estimator [33] is given by:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} \left\{ [z(\mathbf{u}_i) - z(\mathbf{u}_i + \mathbf{h})]^2 \right\}^4 / \left[0,914 + \frac{0,988}{N(\mathbf{h})} \right] \quad (8)$$

where, \mathbf{h} is the vector distance between pairs of observations, $N(\mathbf{h})$ is the number of ordered pairs, and $z(u_i)$ and $z(u_i + \mathbf{h})$ are observed values at their respective locations. The fitting of theoretical semivariogram models was achieved using the Weighted Least Squares (WLS) method, which, according to [34], has the advantage of being able to handle small datasets.

In this work, the Gaussian, Spherical, Exponential, and Wave semivariogram models were used to evaluate which model provides the best spatial prediction results. According to [35] and [36], the models are expressed as:

$$\gamma_{\text{gaus}}(\mathbf{h}) = C_0 + C_1 \left[1 - \exp \left(- \left(\frac{|\mathbf{h}|}{a} \right)^2 \right) \right], |\mathbf{h}| < a, \quad (9)$$

$$\gamma_{\text{exp}}(\mathbf{h}) = C_0 + C_1 \left[1 - \exp \left(- \frac{|\mathbf{h}|}{a} \right) \right], |\mathbf{h}| < a, \quad (10)$$

$$\gamma_{\text{sph}}(\mathbf{h}) = \begin{cases} C_0 + C_1 \left[\frac{3}{2} \left(\frac{|\mathbf{h}|}{a} \right) - \frac{1}{2} \left(\frac{|\mathbf{h}|}{a} \right)^3 \right], |\mathbf{h}| < a, \\ C_0 + C_1, |\mathbf{h}| \geq a \end{cases}, \quad (11)$$

$$\gamma_{\text{wave}}(\mathbf{h}) = \begin{cases} C_0 + C_1 \left[1 - \frac{\sin \left(\pi \frac{h}{a} \right)}{\pi \frac{h}{a}} \right], \\ C_0 + C_1, |\mathbf{h}| \geq a \end{cases}, \quad (12)$$

where, C_0 is the nugget effect and C_1 is the contribution [37]. The range a specifies the distance up to which samples are spatially correlated, and h is the vector distance between pairs of observations.

In this study, both OK and LK methods were employed. LK is variation from Ordinary Kriging, however, with the transformation of the scale of the variable under analysis. The OK estimator for unsampled point is given by:

$$\hat{z}(u_i) = \sum_{i=1}^n \lambda^{OK} z(u_i), \quad (13)$$

where, $\hat{z}(u_i)$ represents the calculated value for location u_i , $z(u_i)$ denotes the observed value, and λ^{OK} corresponds to the weights computed based on n data points.

The LK method is based on transforming the data to the logarithmic scale in order to achieve a symmetric distribution of the data [38]. The transformed data are then used to calculate the semivariogram, and subsequently, OK is employed for interpolating. The obtained values are in the logarithmic domain, necessitating a transformation back to the original scale

$$\hat{z}_{LK}(u_i) = \exp \left[\frac{\hat{y}_{LK}(u_i) + \sigma_{LK}^2(u_i)}{2 - \mu(u_i)} \right], \quad (14)$$

where, $\hat{z}_{LK}(u_i)$ represents the estimates in the original scale, $\hat{y}_{LK}(u_i)$ represents the estimates in the logarithmic scale, σ_{LK}^2 is the variance of LK, and μ is the Lagrange parameter [39].

In order to assess the degree of spatial dependence among maximum rainfall values, the Cambardella criterion (CC) was employed [40], which quantifies the level of spatial dependence. The CC is given by:

$$CC = \frac{C_0}{C_0 + C_1} \times 100, \quad (15)$$

where, C_0 and C_1 are as explained before. The CC is evaluated as follows: when $CC > 75\%$, it can be stated that the variable exhibits weak spatial dependence, for $25\% < CC \leq 75\%$, it is considered to have moderate spatial dependence, and for $CC \leq 25\%$, the variable is deemed to have strong spatial dependence.

In order to evaluate the spatial interpolators of IDW and Kriging, Cross-Validation methods were employed. Three Cross-Validation metrics were used to evaluate the models: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Cross-Validation Mean Prediction Error (CV-MPE). The metrics are defined by the following expressions:

$$MSE = \frac{\sum_{i=1}^n (z(u_i) - \hat{z}(u_i))^2}{n}, \quad (16)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (z(u_i) - \hat{z}(u_i))^2}{n}}, \quad (17)$$

$$CV - MPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{z(u_i) - \hat{z}(u_i)}{z(u_i)} \right| \times 100. \quad (18)$$

In addition to the Cross-Validation metrics, the MPE, as presented in Eq. 6, was also used for evaluating the spatial models. For this particular scenario, MPE was calculated based on the observed and predicted values of maximum rainfall from six locations that were not used for model fitting. The cities in question, along with the station codes, are: Andrelândia (02144019), Bom Jesus da Penha (02146078), Carmo da Cachoeira (02145044), Extrema (02246167), Maria da Fé (02245088) and Santa Rita de Caldas (02246047). An effort was made to select cities where complete sets of daily rainfall data were available for the years 2012 to 2021 since the aim is to compare observed and predicted rainfall for return periods of 2, 5, and 10 years.

2.4 Max-stable processes for analysis of spatial extremes

Considering that the data to be analyzed in this study are extremes, specialized methodologies were employed for the spatial analysis of these variables, namely max-stable processes and the spatial GEV model.

The max-stable process is a class of random fields of significant interest in Extreme Value Theory, as it is suitable for extreme variables [41]. A max-stable process, denoted as $Z(\cdot)$, is the limiting process of maxima of independently and identically distributed random fields with an index set S in the spatial domain, $Y_i(x)_{x \in S}$ with $i = 1, \dots, n$, repetitions of a continuous random process. For two suitable sequences of numbers $a_n(x) > 0$ and $b_n(x) \in \mathbb{R}$, we have that:

$$Z(x) = \lim_{n \rightarrow +\infty} \frac{\max_{i=1}^n Y_i(x) - b_n(x)}{a_n(x)}, \quad x \in S. \quad (19)$$

There are different characterizations of max-stable process models for the analysis of extreme spatial events, such as the Smith model and the Schlatter model. The Smith model for the process $Z(x)$ at locations x_1 and x_2 is given by the bivariate cumulative distribution function [42]:

$$\Pr[Z(x_1) \leq z_1, Z(x_2) \leq z_2] = \exp \left[-\frac{1}{z_1} \Phi \left(\frac{a}{2} + \frac{1}{a} \log \frac{z_2}{z_1} \right) - \frac{1}{z_2} \Phi \left(\frac{a}{2} + \frac{1}{a} \log \frac{z_1}{z_2} \right) \right], \quad (20)$$

where $\Pr[Z(x_1) \leq z_1, Z(x_2) \leq z_2]$ is the bivariate cumulative distribution function, Φ is the standard normal cumulative distribution function defined as $\Phi(u) = (2\pi)^{-1/2} \int_{-\infty}^u \exp(-x^2/2) dx$, and a represents the Mahalanobis distance, defined as $a = \sqrt{(x_1 - x_2)^T \Sigma^{-1} (x_1 - x_2)}$, where Σ is the covariance matrix.

For the Schlather model, the bivariate cumulative distribution function for locations x_1 and x_2 is given by:

$$\Pr[Z(x_1) \leq z_1, Z(x_2) \leq z_2] = \exp \left[-\frac{1}{2} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \left(1 + \sqrt{1 - 2(\rho(\mathbf{h}) + 1) \frac{z_1 z_2}{(z_1 + z_2)^2}} \right) \right], \quad (21)$$

where, $\mathbf{h} \in \mathbb{R}^+$ is the vector of Euclidean distances between two locations.

There are different correlation functions, including Whittle-Matérn, Cauchy, Powered Exponential, and Bessel, given by:

$$\begin{aligned} \text{Whittle-Matérn} & \quad \rho(\mathbf{h}) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\mathbf{h}}{c_2} \right)^\nu K_\nu \left(\frac{\mathbf{h}}{c_2} \right), & c_2 > 0, \nu > 0, \\ \text{Cauchy} & \quad \rho(\mathbf{h}) = \left[1 + \left(\frac{\mathbf{h}}{c_2} \right)^2 \right]^{-\nu}, & c_2 > 0, \nu > 0, \\ \text{Powered Exponential} & \quad \rho(\mathbf{h}) = \exp \left[-\left(\frac{\mathbf{h}}{c_2} \right)^\nu \right], & c_2 > 0, 0 < \nu \leq 2, \\ \text{Bessel} & \quad \rho(\mathbf{h}) = \left(\frac{2c_2}{\mathbf{h}} \right)^\nu \Gamma(\nu + 1) J_\nu \left(\frac{\mathbf{h}}{c_2} \right), & c_2 > 0, \nu \geq \frac{d-2}{2}, \end{aligned} \quad (22)$$

where, c_2 and ν are the range and smoothing parameter, J_ν and K_ν are the Bessel function and the modified Bessel function of the third kind with order ν , and d is the dimension of the random fields.

For a spatial analysis of extremes, it is useful to have knowledge of measures of spatial dependence. One way to assess the degree of spatial dependence is through variograms, however, these may not exist when used for extreme data [43]. Therefore, approaches focused on extremes, such as the Extremal Coefficient and F-madogram, have been proposed as alternatives to overcome this limitation.

The Extremal Coefficient function, aims to assess extreme dependence in max-stable processes by quantifying the dependence between pairs of locations separated by a distance h [44]. The Extremal Coefficient functions $\theta(x_1, x_2)$ for the Smith and Schlather models are given by:

$$\text{Smith: } \theta(x_1, x_2) = 2\Phi(a/2), \quad (23)$$

$$\text{Schlather: } \theta(x_1, x_2) = 1 + \sqrt{\frac{1 - \rho(\|x_1 - x_2\|)}{2}}, \quad (24)$$

where, $1 \leq \theta(x_1, x_2) \leq 2$ with $\theta(x_1, x_2) = 1$ indicating complete dependence between two observations and $\theta(x_1, x_2) = 2$ indicating complete independence.

For a location x with available observations, the GEV distribution can be fitted to the data, and parameter estimates for the distribution can be obtained. However, for a location x without observations, the parameters must be inferred from data at nearby locations. This can be achieved by modeling the spatial evolution of the parameters based on trend surfaces [45].

Trend surfaces can be directly incorporated into the max-stable process by adding extra terms that represent the spatial evolution of GEV parameters. The addition of these terms occurs during model fitting. The trend surfaces adopted for each parameter are:

$$\begin{aligned} \mu(x) &= \beta_{0,\mu} + \beta_{1,\mu} \text{lon}(x) + \beta_{2,\mu} \text{lat}(x) + \beta_{3,\mu} \text{lon}(x) \text{lat}(x) \\ \sigma(x) &= \beta_{0,\sigma} + \beta_{1,\sigma} \text{lon}(x) + \beta_{2,\sigma} \text{lat}(x) + \beta_{3,\sigma} \text{lon}(x) \text{lat}(x), \\ \xi(x) &= \beta_{0,\xi} \end{aligned} \quad (25)$$

where, $\text{lon}(x)$ and $\text{lat}(x)$ represent the longitude and latitude of the locations, and $\beta_{i,\mu}$, $\beta_{i,\sigma}$, and $\beta_{1,\xi}$ are the coefficients of the trend surfaces, with $i = 0, 1, 2, 3$.

The model fitting was performed using the Pairwise Likelihood Maximization method, given by:

$$l_p(z|\boldsymbol{\psi}) = \sum_{k=1}^K \sum_{j=1}^{d-1} \sum_{s=j+1}^d \log f_{j,s}(t(z_{k,j}), t(z_{k,s})|\boldsymbol{\psi}) + \log |J(z_{k,j}) + J(z_{k,s})|, \quad (26)$$

where, $\boldsymbol{\psi}$ is the parameter vector of the model, $f_{j,s}(t(z_{k,j}), t(z_{k,s})|\boldsymbol{\psi})$ is the bivariate density function of the max-stable process, $t(z_{k,j})$ and $t(z_{k,s})$ are the k -th annual maximum observations transformed to Fréchet units for stations j and s , respectively, and $J(z_{k,j})$ and $J(z_{k,s})$ are the Jacobian terms given by: $J(z_{k,j}) = \frac{1}{\sigma(x)} \left(1 + \xi(x) \frac{z_{k,j} - \mu(x)}{\sigma(x)} \right)^{-\frac{1}{\xi(x)}}$.

According to [43], since trend surfaces are used, it is more convenient to omit spatial dependence for calculating the return level prediction for an unobserved location. In this way, it is assumed that the meteorological stations are mutually independent.

Despite the fact that this strategy does not take into account all the uncertainties of the max-stable process parameters, it is expected to be effective, as spatial dependence parameters and trend surface parameters are approximately orthogonal [46].

In addition to the max-stable models, the spatial GEV model was used for the analysis. In this model, it is assumed that observations at different locations are mutually independent [47]. Therefore, the spatial GEV model is given by:

$$Z(x) \sim GEV(\mu(x), \sigma(x), \xi(x)), \quad (27)$$

where, $\mu(x)$, $\sigma(x)$, and $\xi(x)$ are trend surfaces of the form presented in Eq. 25.

The estimation of the parameters of the spatial GEV model can be performed using the Maximum Likelihood Estimation (MLE) technique. According to [46], the log-likelihood of the spatial GEV model is given by:

$$l(z_j(x_i); \boldsymbol{\beta}) = \sum_{i=1}^m \sum_{j=1}^n \left\{ -\log(\sigma(x_i)) - \left(1 + \xi(x_i) \frac{z_j(x_i) - \mu(x_i)}{\sigma(x_i)} \right)^{-\frac{1}{\xi(x_i)}} - \left(1 + \frac{1}{\xi(x_i)} \right) \log \left(1 + \xi(x_i) \frac{z_j(x_i) - \mu(x_i)}{\sigma(x_i)} \right) \right\}, \quad (28)$$

where, $\boldsymbol{\beta}$, $\mu(x_i)$, $\sigma(x_i)$, and $\xi(x_i)$ are the parameters of the spatial GEV model, with x_i representing the i -th location for $i = 1, 2, \dots, m$, and $z_j(x_i)$ representing the j -th observation for the i -th location, with $j = 1, 2, \dots, n$.

According to Gaume [45], based on the fitting of the max-stable process and the spatial GEV model, the return level ($z_T(x)$) can be calculated for an unsampled location x and a return period T using the expression:

$$z_T(x) = \hat{\mu}(x) - \frac{\hat{\sigma}(x)}{\hat{\xi}(x)} \left\{ 1 - \left[-\ln \left(1 - \frac{1}{T} \right) \right]^{-\hat{\xi}(x)} \right\}, \quad (29)$$

where, the estimates of the trend surfaces $\hat{\mu}(x)$, $\hat{\sigma}(x)$, and $\hat{\xi}(x)$ for the location x are inserted.

For the evaluation of max-stable models fit, the Takeuchi's Information Criteria (TIC), proposed by [48], was used, which is a generalization of Akaike's information criterion. According to this criterion, the best-fitting model is considered to be the one that yields the lowest TIC value.

According to [49], the TIC function is given by:

$$TIC = -2l_p(\hat{\boldsymbol{\psi}}) + 2tr\{H(\hat{\boldsymbol{\psi}})^{-1}J(\hat{\boldsymbol{\psi}})\}, \quad (30)$$

where, $\hat{\boldsymbol{\psi}}$ is the vector of parameters of the fitted model being evaluated.

After selecting the most suitable models based on the TIC results, the spatial predictions of these models were evaluated by calculating the MPE (6) of maximum rainfall for the locations of meteorological stations in the municipalities

of Andrelândia, Bom Jesus da Penha, Carmo da Cachoeira, Extrema, Maria da Fé, and Santa Rita de Caldas, for return periods of 2, 5, and 10 years.

Therefore, after obtaining the most suitable spatial extreme model for assessing maximum rainfall in the region, the model was fitted to complete series of extremes with the goal of performing spatial prediction for return periods of 50 and 100 years in the southern and southwestern regions of Minas Gerais. The same procedure was applied to the IDW and Kriging models.

All statistical analyses were conducted using the R software [50].

3 Results and discussion

3.1 Analysing spatial maximum rainfall data through Bayesian-Geostatistical method

The descriptive statistics of the maximum rainfall data, covering the years 1980 to 2021, for the 29 cities in the south and southwest of Minas Gerais analyzed in this study, are presented in Table 2. The results showed that the city of Silvianópolis recorded the highest maximum rainfall, with a value of 226.3 mm. The municipality of Caxambu exhibited the highest variability in the data, with a coefficient of variation equal to 48.23

In Bocaina de Minas, the maximum rainfall data exhibited the highest means and medians, with values of 87.6 mm and 83.5 mm, respectively. These values contrast with those observed in Pouso Alegre, which recorded the lowest means (66.6 mm) and medians (62.1 mm) of maximum rainfall.

Regarding the values of the means and medians of maximum rainfall, it can be observed that for all locations, the average value of rainfall was higher than the median, suggesting that the empirical distributions of the dataset for all municipalities are right-skewed. It can be observed that the Ljung-Box test results for independence indicate that all maximum series are independent at the 5% significance level, as the p -values for all cases were greater than the adopted significance level. Analyzing the Mann-Kendall test results, there is no evidence of trends in the series at the 5% significance level. According to the Kolmogorov-Smirnov test results, the GEV distribution fits the maximum rainfall data of the analyzed cities (Table 3).

Once the assumptions that the maximum series of the analyzed locations are independent and that there is no evidence of trends are met, it is possible to fit the GEV distribution through Bayesian Inference to the maximum rainfall data of each municipality.

In order to verify the convergence of the posterior chains of the GEV distribution parameters, the Geweke, Raftert-Lewis and Heidelberger-Welch criteria were used. The results obtained showed that for all prior structures and all parameters (μ, σ, ξ), there is no evidence of non-convergence of the posterior chains, since, by analyzing the dependence factor

Table 2 Descriptive statistics of annual maximum rainfall (mm) that occurred from 1980 to 2021 in each locality

Cities	\bar{X}	M_d	Max	CV	Cities	\bar{X}	M_d	Max	CV
Alfenas	80.4	79.9	127.9	29.88	Baependi	71.1	69.3	123.0	26.84
Bocaina de M.	87.6	83.5	150.7	29.72	Brazópolis	72.1	68.1	146.5	32.30
Camanducaia	72.8	68.9	200.6	43.32	Cambuí	69.9	65.6	121.6	33.98
Cambuquira	78.8	73.1	161.0	38.01	Campos G.	81.3	78.0	143.0	26.12
Carvalhos	81.1	79.3	167.3	38.79	Caxambu	71.2	70.0	226.0	48.23
Consolação	71.4	67.6	137.8	33.52	Cruzília	68.7	68.2	130.2	28.91
Delfim M.	69.5	60.5	219.3	46.05	Delfinópolis	75.5	68.6	197.2	31.28
Guaxupé	73.4	68.9	120.2	28.28	Illicínea	76.3	71.8	135.4	28.94
Itajuba	69.2	68.6	141.3	28.59	Jacutinga	76.6	72.7	166.9	30.44
Machado	79.6	79.0	126.5	24.02	Muzambinho	74.5	71.5	149.0	29.33
Poço Fundo	79.6	75.9	135.4	25.74	Poços de C.	72.4	71.3	166.6	34.22
Pouso Alegre	66.6	62.1	104.3	28.06	Pouso Alto	74.9	71.0	160.8	33.61
Santa Rita do S.	73.4	68.0	185.2	40.26	São Lourenço	79.1	78.4	171.0	31.25
Silvianópolis	75.5	63.7	226.3	40.45	Três Corações	80.1	78.2	128.0	27.66
Três Pontas	78.9	76.3	131.3	25.03					

Note: Coefficient of variation (CV); Mean (\bar{X}); Median (M_d); Maximum (Max)

Table 3 Results of tests (p -value), Ljung-Box (L-B), Mann-Kendall (M-K), and Kolmogorov-Smirnov (K-S) at a significance level of 5%, regarding annual maximum rainfall data (mm)

Cities	L-B	M-K	K-S	Cities	L-B	M-K	K-S
Alfenas	0.2711	0.2083	0.0603	Baependi	0.5859	0.0564	0.0630
Bocaina de M.	0.7018	0.3918	0.0573	Brazópolis	0.6845	0.2207	0.0835
Camanducaia	0.9969	0.2551	0.1148	CambuÍ	0.7568	0.9421	0.1010
Cambuquira	0.6511	0.3392	0.0699	Campos G.	0.6511	0.3392	0.0699
Carvalhos	0.5007	0.0530	0.0783	Caxambu	0.3029	0.2962	0.1374
Consolação	0.1345	0.0548	0.1067	Cruzília	0.2466	0.0545	0.0660
Delfim M.	0.1067	0.0603	0.0717	Delfinópolis	0.4421	1.0000	0.0746
Guaxupé	0.8015	0.2418	0.0593	Ilicínea	0.6511	0.3392	0.0699
Itajuba	0.2143	1.0000	0.0739	Jacutinga	0.4233	0.3626	0.0538
Machado	0.2546	0.9479	0.0751	Muzambinho	0.5670	0.7615	0.0747
Poço Fundo	0.9568	0.5225	0.0726	Poços de C.	0.4518	0.4873	0.0949
Pouso Alegre	0.9204	0.3615	0.1284	Pouso Alto	0.7701	0.2962	0.0531
Santa Rita do S.	0.6854	0.1500	0.0512	São Lourenço	0.0923	0.1997	0.0687
Silvanópolis	0.5662	0.4350	0.1725	Três Corações	0.4817	0.0596	0.0790
Três Pontas	0.5110	0.0770	0.8485				

Table 4 Mean Prediction Error (MPE %) results for annual maximum rainfall (mm) predicted for return periods of 2, 5, and 10 years, for each location, through Bayesian Inference using non-informative prior (NIP) and informative prior (IP)

Cities	NIP	IP	Cities	NIP	IP
Alfenas	11.40	10.08	Baependi	6.68	9.37
Bocaina de M.	28.22	30.07	Brazópolis	12.98	14.38
Camanducaia	19.35	14.39	CambuÍ	9.46	19.23
Cambuquira	17.19	23.31	Campos G.	22.43	23.01
Carvalhos	39.60	41.16	Caxambu	33.04	38.91
Consolação	48.60	43.79	Cruzília	13.23	14.17
Delfim M.	45.29	33.17	Delfinópolis	12.07	11.29
Guaxupé	21.43	23.41	Ilicínea	8.75	8.87
Itajuba	17.65	14.07	Jacutinga	9.46	3.80
Machado	12.07	16.42	Muzambinho	9.14	9.51
Poço Fundo	4.20	6.35	Poços de C.	28.19	26.17
Pouso Alegre	13.86	10.13	Pouso Alto	6.55	7.80
Santa Rita do S.	13.77	11.28	São Lourenço	16.19	20.45
Silvanópolis	18.48	8.95	Três Corações	14.67	20.47
Três Pontas	14.42	10.00			

of the Raftery-Lewis criterion, it was verified that the values are close to 1, which indicates independence between the iterations. Using the Geweke criterion, it was found that the absolute values obtained were less than 1.96, indicating that there is no evidence of lack of convergence. By means of the Heidelberger and Welch criterion, it was found that their p -values were higher than the significance level (5%), therefore, it can be stated that the chains were stationary.

The results of the convergence criteria of the posterior chains are not presented in the text due to the large amount of information that would be necessary to present, since each of the three tests were applied to the chains of the three GEV parameters, and two models (non-informative and informative) were used for each of the 29 locations.

Once it was verified that there is no evidence of non-convergence of the posterior chains, it was possible to calculate the levels of maximum rainfall return for the 29 locations and for return periods of 2, 5, and 10 years through Bayesian Inference. By obtaining the rainfall estimates, it was possible to evaluate the best prior structure for each city through the calculation of the MPE (Table 4).

It is observed in Table 4 that among all presented values of MPE, the lowest was obtained for the estimates of maximum rainfall in the city of Poço Fundo, with an MPE of 4.20%. This MPE value is considered low when compared to other works that evaluated rainfall estimates for different return periods. For example, [51] analyzed the annual maximum daily rainfall in the city of Silvanópolis-MG using the GEV distribution fitted by Bayesian Inference with different *prior* structures

Table 5 Estimation of the best semivariogram models for each return period (RP) with their respective Cambardella criterion (CC) values and a comparison between the Inverse Distance Weighting (IDW) method using Cross-Validation results: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Cross-Validation Mean Prediction Error (CV-MPE), as well as Mean Prediction Error (MPE) of the six locations not used for model fitting

RP	Krig	Model	C_0	C_1	a	CC	Cross-validation			MPE
							MSE	RMSE	CV-MPE	
2	–	IDW	–	–	–	–	14.71	3.83	4.18	15.29
	LK	Gaus	0.0003	0.0076	0.3925	4.16	12.45	3.52	3.95	14.86
5	–	IDW	–	–	–	–	36.06	6.01	4.90	13.19
	OK	Sph	0.9557	60.3665	0.5785	1.55	32.76	5.72	4.57	13.09
10	–	IDW	–	–	–	–	91.82	9.58	6.39	10.90
	OK	Wave	36.4053	56.4113	0.3001	39.22	78.50	8.86	6.16	9.24

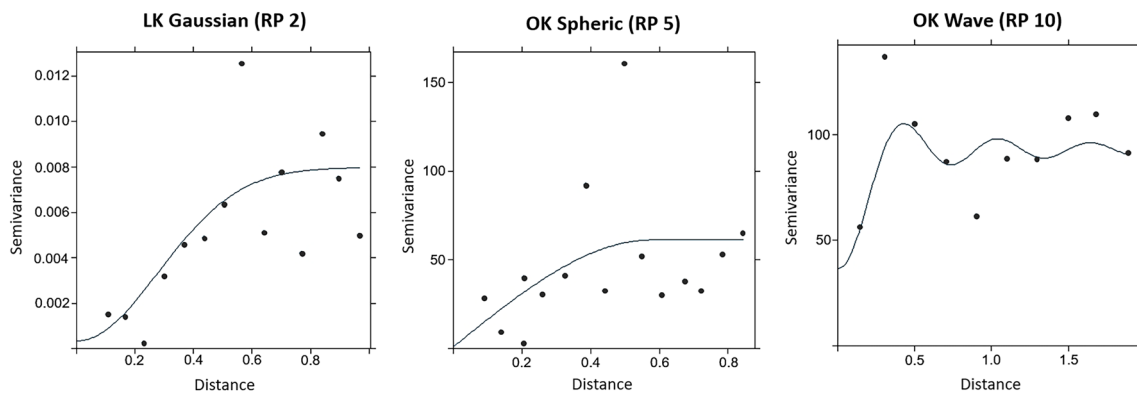


Fig. 2 Graphs of the semivariograms adjusted through weighted least squares method to the empirical semivariograms referring to maximum rainfall data for the return periods of 2, 5, and 10 years, respectively

and through maximum likelihood, and obtained a minimum MPE of 21.53%. Meanwhile, [52] found the lowest MPE of maximum rainfall for dry and rainy periods equal to 16.44%, when analyzing the rainfall of the cities of Machado, São Lourenço, and Juiz de Fora-MG using the GEV, Gumbel and Log-Normal distributions.

Based on the MPE results presented in Table 4, it is possible to observe that by using informative priors based on data from the nearest neighboring city of each locality, better estimates of maximum rainfall were obtained in 12 of the 29 cities. On the other hand, by using non-informative *prior* information, better estimates were obtained in 17 cities. Thus, the maximum rainfall estimates for the 2, 5, and 10-year return periods obtained through Bayesian Inference using the most suitable prior structure for each locality were used.

With possession of the predicted rainfall values, as well as the geographic coordinates of each meteorological station, it is possible to spatialize the attribute of interest over the study area, in this case, the southern and southwestern regions of Minas Gerais. For spatialization through IDW, only the predicted rainfalls and their respective locations are necessary. However, for interpolation through Kriging, it is necessary to fit the semivariograms.

For the calculation and fit of the empirical semivariogram, the robust Cressie-Hawkins semivariogram was used, and the adjustment was performed through the weighted least squares method.

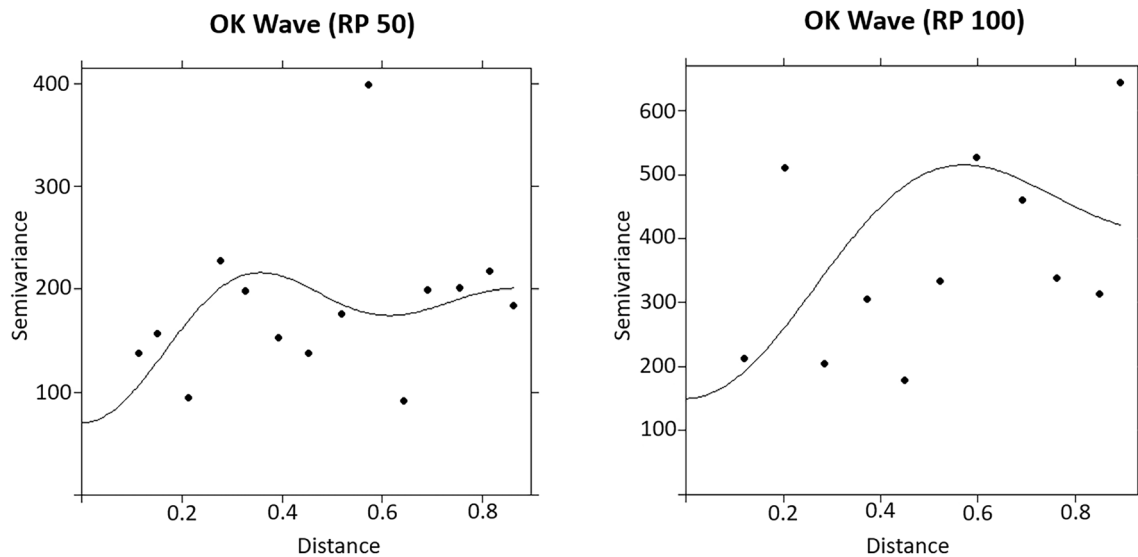
It should be highlighted that in this study, 24 different Kriging methods were used, as two Kriging methods (OK and LK) were associated with four semivariogram models (Gaussian, Exponential, Spherical, and Wave), fitted to the predictions obtained through Bayesian Inference for three return periods. Thus, only the methods that presented the best results for each return period were presented in this section.

The results of the best fits of the Kriging methods associated with their respective best semivariograms, along with the results obtained through IDW, for each return period can be found in Table 5 and Fig. 2. The models evaluation was performed through cross-validation and the calculation of the MPE.

The Kriging method that provided the best spatial predictions for the 2-year return period was the LK with a Gaussian semivariogram model. For the 5- and 10-year return periods, OK was the most suitable for adjustment, using the

Table 6 Estimation of the parameters of Wave semivariograms for rainfall predictions obtained through Bayesian Inference for return periods (RP) of 50 and 100 years

RP	Kriging	Model	C_0	C_1	a
50	OK	Wave	70.0	120.0	0.25
100	OK	Wave	150.0	300.0	0.40

**Fig. 3** Wave semivariogram plots for the empirical semivariograms of rainfall predictions obtained through Bayesian Inference for 50 and 100 years return periods

respective semivariogram models: Spherical and Wave. Analyzing the results of the degree of spatial dependence, it is observed that the Gaussian (RP=2) and Spherical (RP=5) semivariograms showed strong spatial dependence, since the CC results are below 25%. On the other hand, the Wave semivariogram for the 10-year return period showed moderate spatial dependence ($25\% < CC \leq 75\%$).

The cross-validation and MPE results calculated from the IDW interpolation were inferior to the Kriging methods for the three return periods analyzed (Table 5). This result is consistent with that obtained by [53], who found that Geostatistical methods performed better compared to IDW for analyzing monthly and annual rainfall in northeast of Iran. However, it should be noted that except for the mean square error related to the 10-year return period data, all other evaluation criteria presented relatively close results. This low difference between comparisons of interpolation methods was also observed by [54], who analyzed rainfall in the Federal District of Brazil using different interpolation methods, including OK and IDW.

Based on the results, we noticed to perform spatial interpolation of the Bayesian predictions for the 50- and 100-year return periods using OK with the Wave semivariogram model. Thus, the Wave semivariograms were fitted to the maximum rainfall predictions for each location for the return periods of 50 and 100 years obtained through Bayesian Inference. The results of the fitting process are presented in Table 6 and Fig. 3. The prediction maps for the respectively return periods are presented in Fig. 4.

In the prediction map for the 100-year return period, it is observed that more intense rainfall is expected in the extreme north, center, and southeast of the study region. Additionally, higher rainfall is expected in the vicinity of the city of Poços de Caldas.

In the prediction map for the 50-year return period, minimum values of expected maximum rainfall are observed near 118 mm, while the highest values on the map are around 163 mm. As for the prediction map for the 100-year return period, the lowest values of expected maximum rainfall are near 120 mm, and the highest values are around 180 mm.

Although Kriging techniques are often recommended for spatial interpolation, it is important to highlight that they do not always yield the best results [55]. In fact, comparative studies that have employed different interpolation methods, including Kriging, for the analysis of spatial rainfall have shown different models to be more suitable for different

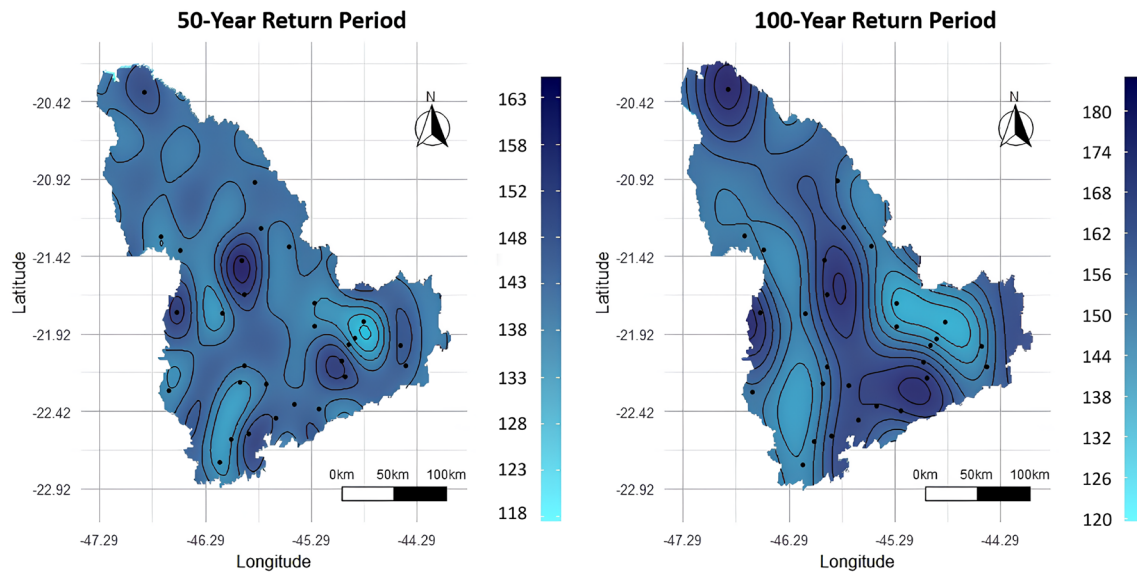


Fig. 4 Maps of predictions of maximum annual rainfall (mm) for the return periods of 50 and 100 years, obtained through Ordinary Kriging with the Wave semivariogram model

study areas. In a study by [56], which compared the performance of seven interpolation models in the analysis of annual mean rainfall in Portugal, it was found that the Empirical Bayesian Kriging with Regression model proved to be the most appropriate for the analysis. The other models used in the study included Local Polynomial Interpolation, Global Polynomial Interpolation, Radial Basis Function, IDW, Ordinary Cokriging, and Universal Cokriging. In the study conducted by [57], six interpolation models were evaluated for the analysis of annual mean rainfall in Ontario, Canada. The analyzed methods included IDW, Global Polynomial Interpolation, Local Polynomial Interpolation, Radial Basis Function, OK, and Universal Kriging. The results indicated that the most suitable model for the analysis of mean rainfall in the region was Local Polynomial Interpolation. Reference [58] investigated the annual maximum rainfall in the Haihe River basin, located in northern China, using different interpolation methods to assess their performance. The methods employed for the analysis included IDW, OK, External Drift Kriging with different covariates, and Empirical Bayesian Kriging. The results showed that the External Drift Kriging method, incorporating the annual mean rainfall as a covariate, yielded the best results among the evaluated interpolation methods.

3.2 Max-stable process and spatial GEV

Unlike the approach through Kriging and IDW employed in this study, which use only one observation for each location for spatial interpolation, in the spatial analysis of extremes using max-stable processes and spatial GEV, each location in the region is associated with a series of annual maximum rainfall values.

To fit the max-stable models via Pairwise Likelihood Maximization and assess spatial dependence, the first step is to transform the data from the maximum series to Fréchet units. To achieve this, the GEV distribution was fitted to the rainfall data from each location using the Maximum Likelihood method, aiming to obtain estimates for each parameter and thus transform the data to the new scale.

Thus, it was possible to fit the max-stable models, obtaining estimates of the dependency parameters for the Smith and Schlather models. The model fit was also assessed using the TIC. In our study, the best results were found across the Smith model and Schlather model with Bessel correlation, and the spatial GEV model.

Through the fitting of max-stable models, it is possible to obtain estimates of the dependency parameters as well as to estimate the coefficients β of the trend surfaces. As for the spatial GEV model, from its fitting, it is possible to obtain estimates of its parameters, which include β and the trend surfaces (Table 7).

To assess the spatial dependence structure of extremes using the max-stable models, the Extremal Coefficient was used. The evaluation of spatial dependence in extremes is done visually through a graph of the Extremal Coefficient as

Table 7 Estimation of the coefficients of the trend surfaces (t.s.) for the parameters of the GEV distribution

Models	t.s.*	β_0	β_1	β_2	β_3
Smith	$\mu(x)$	2547.19	105.11	51.66	2.18
	$\sigma(x)$	1652.62	72.14	35.92	1.58
	$\xi(x)$	-0.02	-	-	-
Schlather (Bessel)	$\mu(x)$	2547.20	104.81	51.74	2.17
	$\sigma(x)$	1652.69	72.07	35.89	1.58
	$\xi(x)$	-0.01	-	-	-
spatial GEV	$\mu(x)$	2547.19	105.22	51.62	2.18
	$\sigma(x)$	1632.23	72.07	35.84	1.60
	$\xi(x)$	-0.01	-	-	-

*The models for $\mu(x)$, $\sigma(x)$ and $\xi(x)$ are the given in Eq. 25

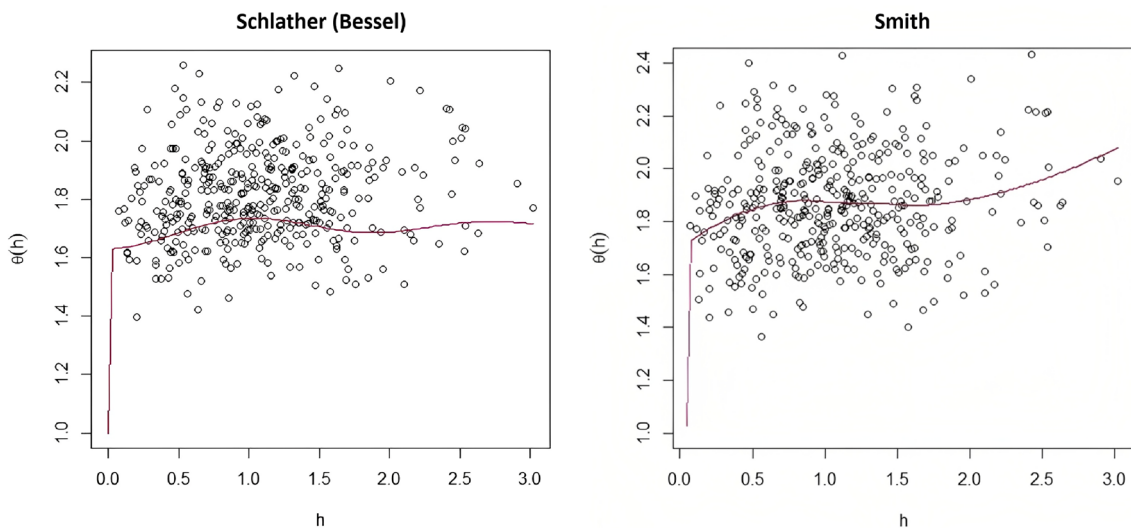


Fig. 5 Extremal coefficient for the Smith and Schlather max-stable models with Bessel correlation function, related to the maximum rainfall data from 1980 to 2011

a function of the distance between pairs of observations. The Extremal Coefficient graph for the Smith and Schlather (Bessel) max-stable models in relation to pairs of meteorological stations is presented in Fig. 5.

Analyzing the Extremal Coefficient graphs presented in Fig. 5, it can be observed that for the Smith model, spatial dependence decreases as the distance between stations increases. This occurs because the closer $\theta(h)$ is to 2, the weaker the dependence between locations. In addition, for both models, the Extremal Coefficient presents values greater than 1.4 regardless of the distance h . According to [59], who obtained similar results in the analysis of spatial dependence, an Extremal Coefficient value greater than 1.4 suggests that the dependence between pairs of observations is low, indicating low evidence of dependence for the annual maximum rainfall data.

Table 8 Results of the Mean Prediction Error (MPE), given in percentage, for predictions at the six locations that were not used in model fitting, obtained from the Smith and Schlather max-stable models with Bessel correlation, in addition to the spatial GEV model

Model	MPE			$\bar{X}(MPE)$
	TR 2	TR 5	TR 10	
Smith	15.17	13.72	10.98	13.29
Schlather (Bessel)	15.64	14.03	14.12	14.60
GEV espacial	15.41	13.75	11.90	13.69

Mean of the Mean Prediction Error ($\bar{X}(MPE)$)

Table 9 Results of fitting the Smith max-stable model to the maximum rainfall data for the period from 1980 to 2021

t.s.	β_0	β_1	β_2
$\mu(x)$	3468.7	150.8	71.2
$\sigma(x)$	2004.3	87.5	43.4
$\xi(x)$	- 0.03	-	-
d.p.	σ_{11}	σ_{12}	σ_2
-	0.2386	- 0.4277	0.7669

Trend surface (t.s.); Dependency parameter (d.p.)

According to [59], one way to improve the assessment of spatial dependence would be to use the approach of selecting extremes above a threshold. This approach would allow for better utilization of available data, thereby increasing the amount of data to be used for model fitting.

Based on the results regarding the spatial dependence of extremes, omitting the dependence for spatial prediction calculations for different return periods, as previously discussed, becomes a plausible strategy. In this way, spatial rainfall predictions were calculated for return periods of 2, 5, and 10 years for the southern and southwestern regions of Minas Gerais.

By the fitted models, spatial predictions can be calculated, which allowed for model evaluation through the calculation of the MPE for rainfall at the six locations that were not used in model fitting. The results of the MPE are presented in Table 8.

It is possible to observe that the three models demonstrated similar performances in spatial predictions for return periods of 2 and 5 years, with subtle differences in the MPE values (Table 8). Despite the proximity of the results, it was found that using the Smith model yielded the lowest MPE values for all the considered return periods. Therefore, it was resolved to make spatial predictions for return periods of 50 and 100 years in the southern and southwestern regions of Minas Gerais using the Smith max-stable model. The results of fitting the Smith model, now considering the complete maximum rainfall series from 1980 to 2021, is presented in Table 9.

Through fitting the Smith max-stable model and subsequently obtaining the coefficients of the trend surfaces, spatial predictions were calculated, and maps of maximum rainfall prediction for the southern and southwestern regions of Minas Gerais for return periods of 50 and 100 years were obtained. These maps are presented in Fig. 6.

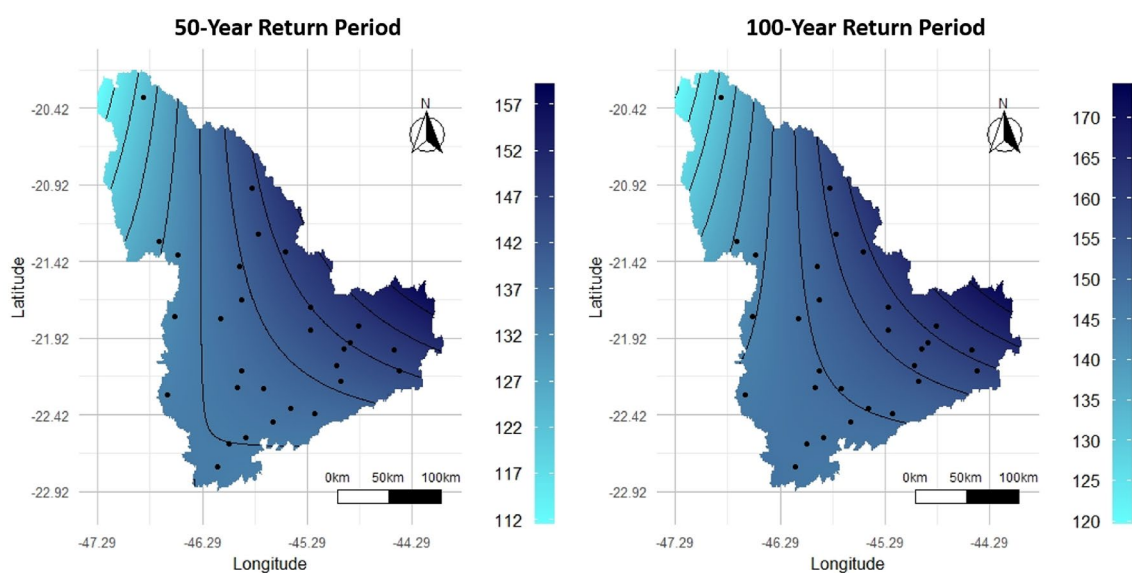


Fig. 6 Maps of maximum rainfall prediction (mm) for return periods of 50 and 100 years, obtained using the Smith max-stable model

When analyzing the prediction maps for return periods of 50 and 100 years, it can be observed that more intense rainfall events are expected in the eastern portion of the analyzed region, while areas located to the north are expected to have lower volumes of daily maximum rainfall.

However, it is noted that the spatial distribution of the expected maximum rainfall for the two return periods is similar, with significant differences primarily in the scale of expected rainfall. Furthermore, the maps show that the spatial distribution of rainfall occurs continuously and smoothly in the region, without the formation of isolated contours.

This characteristic was also observed in the work of [60], who analyzed maximum rainfall in Switzerland using Copula models and max-stable processes. According to the authors, a disadvantage of max-stable processes is the difficulty in finding suitable trend surfaces, which can result in unrealistically smooth prediction maps.

4 Conclusion

The maximum rainfall series from the 29 analyzed cities are independent and show no evidence of trend. The Generalized Extreme Value distribution provided a good fit to the maximum series of the 29 locations.

The results of the Mean Prediction Error for maximum rainfall obtained through Bayesian Inference showed that the use of non-informative priors led to better prediction results for 17 municipalities. On the other hand, the use of informative priors based on data from the nearest location yielded better results for 12 locations.

The Gaussian and Spherical semivariograms exhibited a strong degree of spatial dependence for the data corresponding to return periods of 2 and 5 years, respectively. Meanwhile, the Wave semivariogram, fitted to the rainfall data for a return period of 10 years, showed a moderate spatial dependence.

The models that provided the best spatial predictions for return periods of 2, 5, and 10 years, among the Kriging and Inverse Distance Weighting methods, were the Log-Normal Kriging with a Gaussian semivariogram model, Ordinary Kriging with a Spherical semivariogram model, and Ordinary Kriging with a Wave semivariogram model, respectively.

For the 50-year return period, it is expected that more intense rainfall will occur in the center and surrounding areas of the city of Poços de Caldas. As for the 100-year return period, higher daily rainfall volumes are expected in the extreme north, center, and southeast of the region, as well as in Poços de Caldas neighborhood.

For the evaluation of spatial extremes models, it was found that the Smith and Schlather max-stable models with Bessel correlation outperform. Through the Mean Prediction Error calculation, it was observed that the models showed similar results, however, the Smith model proved to be more suitable for the analysis of maximum rainfall in the southern and southwestern regions of Minas Gerais.

The fitted Extremal Coefficient for the Smith and Schlather max-stable models with Bessel correlation showed low evidence of spatial dependence in the maximum rainfall data.

The prediction maps for return periods of 50 and 100 years obtained from the Smith max-stable model exhibited a similar spatial variability structure for rainfall. Furthermore, the spatial variation occurred continuously and smoothly throughout the analyzed region.

It is worth noting that both methodologies used in this study constitute viable analysis alternatives. Using the method discussed in section 3.1, it was found that the best geostatistical approach using Bayesian predictions presented mean average prediction error similar to that found with spatial gev. Therefore, when it comes to using a certain methodology, time and expertise can be taken into consideration to obtain spatial predictions, that is, in the Bayesian approach, convergence of the Markov chains must be preceded for each municipality evaluated, and, in the max-stable approach, one must proceed with the most appropriate model.

Acknowledgements Acknowledgments to INMET and ANA for providing the dataset of this research and to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for providing the master's scholarship which supported this study.

Author contributions T.R.F. (Formal analysis; Investigation; Software; Validation; Writing—original draft) G.R.L.(Conceptualization; Data curation; Formal analysis; Funding acquisition; Methodology; Project administration; Software; Supervision; Validation; Visualization; Writing—review & editing) L.A.B. (Conceptualization; Methodology; Software; Visualization; Writing—review & editing)

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication The authors give the consent to the publisher to publish this work.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Panwar V, Sen S. Economic impact of natural disasters: an empirical re-examination. *Margin J Appl Econ Res.* 2019;13(1):109–39. <https://doi.org/10.1177/0973801018800087>.
2. Viana TV, Alves AM, Sousa VF, Azevedo BM, Furlan RA. Planting density and number of drains influencing the productivity of rose plants cultivated in pots. *Hortic Bras.* 2008;26:528–32. <https://doi.org/10.1590/S0102-05362008000400021>.
3. Nova RIT, Susanna D, Warsito GM. The presence of rodents infected with *Leptospira* bacteria in various countries and the leptospirosis potential in humans: a systematic review. *Malay J Public Health Med.* 2020;20(2):185–96. <https://doi.org/10.37268/mjphm/vol.20/no.2/art.250>.
4. Fisher RA, Tippett LHC. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math Proc Cambridge Philos Soc.* 1928;24(2):180–90. <https://doi.org/10.1017/s0305004100015681>.
5. Stephenson A, Tawn J. Bayesian inference for extremes: accounting for the three extremal types. *Extremes.* 2004;7(4):291–307. <https://doi.org/10.1007/s10687-004-3479-6>.
6. Carmo EJ, Rodrigues DD, Santos GRD. Evaluation of kriging interpolators and topo to raster for the generation of digital elevation models from a “as built”. *Boletim de Ciências Geodésicas.* 2015;21(4):674–90. <https://doi.org/10.1590/S1982-21702015000400039>.
7. Cerón WL, Andreoli RV, Kayano MT, Canchala T, Carvajal-Escobar Y, Souza RAF. Comparison of spatial interpolation methods for annual and seasonal rainfall in two hotspots of biodiversity in South America. *Academia Brasileira de Ciências.* 2021;93(1):1–22. <https://doi.org/10.1590/0001-3765202120190674>.
8. Banerjee S, Carlin BP, Gelfand AE. Hierarchical modeling and analysis for spatial data. New York: Chapman and Hall/CRC; 2014. p. 584. <https://doi.org/10.1201/b17115>.
9. Dantas GD, Oliveira LA. Analysis of spatial continuity of precipitation in the São Francisco river basin in its area of occurrence in the state of Minas Gerais-Brazil, historical series 2004 to 2017. *Braz J Dev.* 2021;7(3):23585–95. <https://doi.org/10.34117/bjdv7n3-190>.
10. Guimarães VL, Alves RC. Comparison of geostatistical models for rainfall forecast in Minas Gerais, Brazil, between 2000 and 2021 hydrological years. *Rev Bras Geografia Física.* 2023;16(1):528–41. <https://doi.org/10.26848/rbgf.v16.1.p528-541>.
11. Oliveira Teixeira SH, Souza AL. Analysis of the geographical distribution of covid-19 in the south/southwestern mesoregion of Minas Gerais. *Hygeia Revista Brasileira de Geografia Médica e da Saúde.* 2020. <https://doi.org/10.14393/hygeia0054632>.
12. INE: Resultados del Censo de Población 2011: Población, Crecimiento Y Estructura Por Sexo y edad. Instituto Nacional de Estadística, <https://www.ine.gub.uy/censos-2011> (2011). Instituto Nacional de Estadística
13. Reboita MS, Rodrigues M, Silva LF, Alves MA. Climate aspects in Minas Gerais state. *Rev Bras Climatol.* 2015;17:206–26. <https://doi.org/10.5380/abclima.v17i0.41493>.
14. Martins FB, Gonzaga G, Dos Santos DF, Reboita MS. Climate classification of Köppen and Thornthwaite for Minas Gerais: current climate and climate changes projections. *Rev Bras Climatol.* 2018;14:129–56. <https://doi.org/10.5380/abclima.v1i0.60896>.
15. Kowarik A, Templ M. Imputation with the R package VIM. *J Stat Softw.* 2016;74(7):1–16. <https://doi.org/10.18637/jss.v074.i07>.
16. Ljung GM, BOX GEP. On a measure of lack of fit in time series models. *Biometrika.* 1978;65(2):297–303. <https://doi.org/10.1093/BIOMET/65.2.297>.
17. Mann HB. Nonparametric tests against trend. *Econometrica.* 1945;13(3):245. <https://doi.org/10.2307/1907187>.
18. Kendall MG. Rank correlation measures. 15th ed. London: Charles Griffin Book Series; 1975. p. 2002.
19. Xie M, Singh K. Confidence distribution, the frequentist distribution estimator of a parameter: a review. *Int Stat Rev.* 2013;81(1):3–39. <https://doi.org/10.1111/insr.12000>.
20. Coles SG, Powell EA. Bayesian methods in extreme value modelling: a review and new developments. *Int Stat Rev.* 1996;64(1):119–36. <https://doi.org/10.2307/1403426>.
21. Oliveira C, Lugon Junior J, Knupp DC, Silva Neto AJ, Prieto-Moreno A, Llanes-Santiago O. Estimation of kinetic parameters in a chromatographic separation model via Bayesian inference. *Rev Int Métodos Numéricos para Cálculo y Diseño en Ingeniería.* 2018;34(1):1–26. <https://doi.org/10.23967/j.rimni.2017.12.002>.

22. Chung E-S, Kim SU. Bayesian rainfall frequency analysis with extreme value using the informative prior distribution. *KSCE J Civ Eng*. 2013;17(6):1502–14. <https://doi.org/10.1007/S12205-013-0189-0>.
23. Nogueira DA, Safadi T, Ferreira DF. Evaluation of univariate convergence criteria for the Monte Carlo method via Markov chains. *Rev Bras Estatística*. 2004;65(224):59–88.
24. Raftery AE, Lewis S. Comment: one long run with diagnostics: implementation strategies for Markov chain. *Stat Sci*. 1992;7(4):493–7. <https://doi.org/10.1214/ss/1177011143>.
25. Geweke J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo J, Berger J, Dawid A, Smith A (eds) *Bayesian statistics 4* Oxford University Press, Oxford; 1992. , pp. 169–193. <https://global.oup.com/academic/product/bayesian-statistics-4-9780198522669?lang=en&cc=gb>
26. Heidelberger P, Welch PD. Simulation run length control in the presence of an initial transient. *Oper Res*. 1983;31(6):1109–44. <https://doi.org/10.1287/opre.31.6.1109>.
27. Martins AP, Santos Alves W, Damasceno CE. Evaluation of interpolation methods for spatialization of air temperature in the Paranaíba River Basin—Brazil. *Rev Bras Climatol*. 2019;25:444–63. <https://doi.org/10.5380/abclima.v25i0.64291>.
28. Achouri M, Gifford GF. Spatial and seasonal variability of field measured infiltration rates on a rangeland site in Utah. *Rangeland Ecol Manag/J Range Manag Arch*. 1984;37(5):451–5. <https://doi.org/10.2307/3899635>.
29. Thompson SK. *Sampling*. 1st ed. New York: Wiley-interscience Publication; 1992. p. 343.
30. Tobin C, Nicotina L, Parlange MB, Berne A, Rinaldo A. Improved interpolation of meteorological forcings for hydrologic applications in a Swiss Alpine region. *J Hydrol*. 2011;401(1):77–89. <https://doi.org/10.1016/J.JHYDROL.2011.02.010>.
31. Batista ML, Coelho G, Reis Teixeira MB, Oliveira MS. Semivariance estimators: analysis of performance in the mapping of annual precipitation. *Sci Agrar*. 2018;19(1):64–74. <https://doi.org/10.5380/rsa.v19i1.53823>.
32. Barbosa DP, Bottega EL, Valente DSM, Santos NT, Guimarães WD. Delineation of homogeneous zones based on geostatistical models robust to outliers. *Revista Caatinga*. 2019;32(2):472–81. <https://doi.org/10.1590/1983-21252019v32n220rc>.
33. Cressie N, Hawkins DM. Robust estimation of the variogram. I. *Math Geol*. 1980;12(2):115–25. <https://doi.org/10.1007/BF01035243>.
34. Carvalho JRP, Vieira SR, Grego CR. Comparison of methods for adjusting semivariogram model of annual rainfall. *Rev Bras Engenharia Agrícola e Ambiental*. 2009;13(4):443–8. <https://doi.org/10.1590/S1415-43662009000400011>.
35. Xavier AC, Cecílio RA, Lima JSS, et al. Matlab modules forspatial interpolation by ordinary kriging and inverse distance. *Rev Bras Cartogr*. 2010;62(1):67–76. <https://doi.org/10.14393/rbcv62n1-43668>.
36. Appel Neto E, Barbosa IC, Seidel EJ, Oliveira M.S.d. Spatial dependence index for cubic, pentaspherical and wave semivariogram models. *Boletim de Ciências Geodésicas*. 2018;24(1):142–51. <https://doi.org/10.1590/S1982-21702018000100010>.
37. Pereira VAS, Pugliesi EA, Flores EF, Camargo PO. Ordinary kriging and depicting uncertainties applied in the monitoring of ionospheric irregularities in Brazil. *Rev Bras Cartogr*. 2018;70(3):967–96. <https://doi.org/10.14393/rbcv70n3-45708>.
38. Webster J, Oliver MA. *Geostat Environ Sci*. Chichester: Wiley; 2007. p. 330. <https://doi.org/10.1002/9780470517277>.
39. Yamamoto JK, Furuie RA. A survey into estimation of lognormal data. *Geociências*. 2010;29(1):5–19.
40. Cambardella CA, Moorman TB, Novak JM, Parkin TB, Karlen DL, Turco RF, Konopka AE. Field-scale variability of soil properties in central Iowa soils. *Soil Sci Soc Am J*. 1994;58(5):1501–11. <https://doi.org/10.2136/SSAJ1994.03615995005800050033X>.
41. Dombry C, Engelke S, Oesting M. Exact simulation of max-stable processes. *Biometrika*. 2016;103(2):303–17. <https://doi.org/10.48550/arXiv.1506.04430>.
42. Thibaud E, Mutzner R, Davison AC. Threshold modeling of extreme spatial rainfall. *Water Resour Res*. 2013;49(8):4633–44. <https://doi.org/10.1002/wrcr.20329>.
43. Ribatet M. Spatial extremes: max-stable processes at work. *J Soc Française Stat*. 2013;154(2):156–77.
44. Azizah S, Sutikno S, Puhadi P. Parameter estimation of smith model max-stable process spatial extreme value (case-study: extreme rainfall modelling in Ngawi Regency). *IPTEK J Sci*. 2017;2(1):16–20. <https://doi.org/10.12962/j23378530.v2i1.a2255>.
45. Gaume J, Eckert N, Chambon G, Naaim M, Bel L. Mapping extreme snowfalls in the French Alps using max-stable processes. *Water Resour Res*. 2013;49(2):1079–98. <https://doi.org/10.1002/wrcr.20083>.
46. Ribatet M. A User's Guide to the SpatialExtremes Package. EPFL, Lausanne, Switzerland (2009). EPFL. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=bf87e33931e3e5ad5e2621cfedc41ca6deb585d8>
47. Cao Y, Li B. Assessing models for estimation and methods for uncertainty quantification for spatial return levels. *Environmetrics*. 2019;30(2):2508. <https://doi.org/10.1002/env.2508>.
48. Takeuchi K. Distribution of an information statistic and the criterion for the optimal model. *Math Sci*. 1976;153:12–8.
49. Senapeng P, Prahadchai T, Guayjarenpnashk P, Park J-S, Busababodhin P. Spatial modeling of extreme temperature in northeast Thailand. *Atmosphere*. 2022;13(4):589. <https://doi.org/10.3390/atmos13040589>.
50. R Core Team: R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2021). R Foundation for Statistical Computing. <https://www.R-project.org/>
51. Martins TB, Almeida GC, Avelar FG, Beijo LA. Prediction of maximum precipitation in the municipality of Silvianópolis-MG: classical and Bayesian approaches. *IRRIGA*. 2018;23(3):467–79. <https://doi.org/10.15809/irriga.2018v23n3p467-479>.
52. Ferreira TR, Beijo LA, Avelar FG. Evaluation of probability distributions in the study of maximum rainfall in three cities in Minas Gerais State. *Rev Bras Climatol*. 2021;29:526–44. <https://doi.org/10.5380/abclima>.
53. Delbari M, Afrasiab P, Jahani S. Spatial interpolation of monthly and annual rainfall in northeast of Iran. *Meteorol Atmos Phys*. 2013;122:103–13. <https://doi.org/10.1007/s00703-013-0273-5>.
54. Borges PA, Franke J, Anunciação YMT, Weiss H, Bernhofer C. Comparison of spatial interpolation methods for the estimation of precipitation distribution in distrito federal, Brazil. *Theoret Appl Climatol*. 2016;123:335–48. <https://doi.org/10.1007/s00704-014-1359-9>.
55. Pereira P, Oliva M, Baltreñaite E. Modelling extreme precipitation in hazardous mountainous areas. contribution to landscape planning and environmental management. *J Environ Eng Landsc Manag*. 2010; 18(4): 329–342. <https://doi.org/10.3846/jeelm.2010.38>
56. Antal A, Guerreiro PM, Cheval S. Comparison of spatial interpolation methods for estimating the precipitation distribution in Portugal. *Theoret Appl Climatol*. 2021;145:1193–206. <https://doi.org/10.1007/s00704-021-03675-0>.

57. Wang S, Huang G, Lin Q, Li Z, Zhang H, Fan Y. Comparison of interpolation methods for estimating spatial distribution of precipitation in Ontario, Canada. *Int J Climatol*. 2014;34(14):3745–51. <https://doi.org/10.1002/joc.3941>.
58. Zou W-Y, Yin S-Q, Wang W-T. Spatial interpolation of the extreme hourly precipitation at different return levels in the Haihe river basin. *J Hydrol*. 2021;598: 126273. <https://doi.org/10.1016/j.jhydrol.2021.126273>.
59. Diriba TA, Debusho LK. Statistical modeling of spatial extremes through max-stable process models: application to extreme rainfall events in South Africa. *J Hydrol Eng*. 2021;26(10): 05021028-1. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0002123](https://doi.org/10.1061/(ASCE)HE.1943-5584.0002123).
60. Davison AC, Padoan SA, Ribatet M. Statistical modeling of spatial extremes. *Stat Sci*. 2012;27(2):161–86. <https://doi.org/10.1214/11-STS376>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.