Assessing intensity-duration-frequency equations and spatialization techniques across the Grande River Basin in the state of Bahia, Brazil

Avaliação de equações de intensidade-duração-frequência e técnicas de espacialização na Bacia do Rio Grande no estado da Bahia, Brasil

Alan de Gois Barbosa[1] [ID], Izaias Rodrigues de Souza Neto[2] [ID], Veber Afonso Figueiredo Costa[1] [ID] & Ludmilson Abritta Mendes[2] [ID]

[1]Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil
[2]Universidade Federal de Sergipe, São Cristóvão, SE, Brasil

E-mails: alanbrb@ufmg.br (AGB), izaiasnetoengcivil@gmail.com (IRSN), veber@ehr.ufmg.br (VAFC), lamendes@academico.ufs.br (LAM)

ABSTRACT

Understanding the probabilistic behavior of extreme rainfall on a fine temporal and spatial scales is crucial for design and risk assessment of hydraulic structures. However, information at appropriate resolutions is frequently limited or unavailable at the locations of interest, thereby requiring the estimation of Intensity-Duration-Frequency (IDF) curves at the regional scale. In this paper, we resort to simplified approaches for rainfall disaggregation and spatialization for deriving a regional IDF equation for the Grande River catchment, in the Brazilian state of Bahia. Our results suggest that, at the daily time scale, the maximum rainfall amounts can be reasonably described by the light-tailed Gumbel distribution in the study region. The spatialization procedures indicated that, whereas both the Inverse Distance Weighting (IDW) and the ordinary kriging techniques could capture the spatial variability of rainfall quantiles, for several durations of practical interest, only the former was able to model the spatial variability of the IDF parameters. Finally, despite the simplifying assumptions, we were able to derive smooth spatial surfaces for the aforementioned quantities, which might be useful for the design of hydraulic structures at ungauged sites.

Keywords: Heavy rainfalls; IDF; IDW; Kriging.

RESUMO

Compreender o comportamento probabilístico de chuvas extremas em escalas temporais e espaciais é crucial para o projeto e avaliação de risco de estruturas hidráulicas. Porém, informações em resolução adequada são frequentemente limitadas ou indisponíveis nos locais de interesse, o que, por sua vez, exigiria a estimativa de curvas de intensidade-duração-frequência (IDF) em escala regional. Neste artigo, empregam-se abordagens simplificadas para desagregação e espacialização de chuvas para derivar uma equação IDF regional para a bacia do Rio Grande, no estado da Bahia. Os resultados sugerem que, na escala de tempo diária, as quantidades máximas de chuva na região de estudo podem ser descritas razoavelmente pela distribuição de Gumbel. Os procedimentos de espacialização indicaram que, enquanto tanto a Ponderação pelo Inverso da Distância (IDW) quanto as técnicas de krigagem ordinária poderiam capturar a variabilidade espacial dos quantis de chuva, apenas a primeira foi capaz de modelar a variabilidade espacial dos parâmetros IDF para várias durações de interesse prático. Finalmente, apesar das premissas simplificadoras, foi possível derivar superfícies espaciais lisas para as grandezas mencionadas, o que pode ser útil para o projeto de estruturas hidráulicas em locais desprovidos de monitoramento.

Palavras-chave: Chuvas intensas; IDF; IDW; Krigagem.

INTRODUCTION

Reliable rainfall estimates, at suitable temporal and spatial scales, are required for planning and designing water infrastructure for flood conveyance, and for developing strategies for flood hazard mitigation. In this sense, the probabilistic understanding on extreme rainfall at a given site or region is necessary for supporting decision making processes regarding the cost-risk tradeoff in such projects (Nunes et al., 2021). In general, for the sub daily time scales, the random behavior of extreme rainfall is summarized by Intensity-Duration-Frequency (IDF) models, which accommodate, in a single parametric form, the functional relationships among the maximum rainfall intensities and their durations and frequencies (or return periods) (Sabino et al., 2020).

Extreme rainfall events are increasing around the world and in Brazil, originating large floods more frequently, decreasing the economic productivity and causing damages and economic losses in both urban and rural areas (Aragão et al., 2013; Asadieh & Krakauer, 2017). Large rainfall datasets can be used to understand the random behavior of rainfall, so the study of the variability of this variable and its spatiotemporal distribution can help handling changes in patterns of precipitation at several scales of interest (Bougara et al., 2020).

Using uninterrupted records of rainfall data from rain gauges is the most usual alternative to estimate an IDF relationships (Souza et al., 2019). However, in many regions of the world there is a reduced number of pluviograph information, which hinders the probabilistic modeling of extreme rainfall at sub daily time scale. Then, disaggregation techniques for estimating short duration rainfall are useful for dealing with this problem. Parametric and nonparametric approaches que been extensively utilized in this context. The former might encompass both point processes models and the self-similarity (or fractal) approach (Diez-Sierra & del Jesus, 2019) and require sub daily information at the target site, which is, more often than not, limited or unavailable. The latter, in turn, resorts to resampling techniques for deriving sub daily rainfall amounts from the data at the target sites themselves or from neighbor gauges (e.g., Aguilar & Costa, 2020), but are computationally expensive and have limited extrapolation abilities.

An alternative to aforementioned difficulties is using disaggregation indices in daily rainfall data series, i.e., from pluviometric stations (Back, 2020). In Brazil the most used method for disaggregation is that by CETESB (Aragão et al., 2013), which, albeit not based on a rigorous mathematical formalism, relies, to some extent, on the self-similarity rationale (Serinaldi, 2010) for linking maximum rainfall intensities across short time scales over large regions.

In addition to the usual scarcity of rainfall information at fine time resolution, many regions in Brazil still lack IDF relationships due to the limited density of rainfall gauging stations with either daily or sub daily records. An example is the western region of the state of Bahia, in which there is a growing demand for water to supply farmers and irrigation projects in several watersheds. The Grande River Basin is one of the most important contributors to the São Francisco flow rates, having the highest potential contribution (14.2%), and the second highest real contribution (10.9%) (Pereira et al., 2007). Many rivers in Grande River catchment have already reached the legal limits for granting water use, which has led to serious conflicts among users (Gonçalves et al., 2020).

Also, according to Jesus and Nascimento (2020), the region of the Rio Grande catchment is frequently affected by extreme rainfall events, as the state of Bahia is exposed to different meteorological systems, but the limited availability of rainfall data might hinder the appropriate management of water resources (Moreira & Silva, 2010), mainly with respect to flood risk assessment and flood mitigation. Up to 2020, the municipality of Barreiras was the only city with a properly defined IDF relationship in the Grande River catchment - although 15 equations have been recently determined for the region by Moreira et al. (2020) on the basis of relatively small datasets and potentially high levels of estimation uncertainty.

The use of geostatistical tools for spatialization of extreme rainfall quantiles constitutes an alternative for regions with scarcity of rain gauges. In these cases, the interpolation of quantiles associated to pre-established durations frequently used for hydraulic design may provide predictions of the quantities of interest (Souza et al., 2019). The main techniques for this purpose are the Inverse Distance Weighting (IDW) and the Ordinary Kriging (Pizarro et al., 2018; Rabelo et al., 2018; Souza et al., 2019), the former being based on a deterministic approach whereas the latter stems from the theory of geostatistics (Ly et al., 2013).

In view of the foregoing, the objectives of this study are to determine IDF relationships for the region of the Grande River catchment, and to elaborate isoline maps for the parameters of the IDF equations and for quantiles usually utilized in urban drainage design. The results of the study might provide a better understanding of the spatial distribution of extreme precipitation across the region, as well as provide quantile estimates for design purposes. The remainder of the paper is organized as follows. In Section 2, a brief description of the study region and the methodological steps are provided, with focus on the theoretical reasoning and the simplifying assumptions utilized for deriving our models. Section 3 presents the results for at-site IDF modeling and the subsequent spatial analyses for estimating IDF parameters and quantiles at ungauged locations. Finally, in Section 4, we present the concluding remarks of the study and future research developments.

MATERIAL AND METHODS

The Grande River catchment is located in the Brazilian state of Bahia (Figure 1), with geographic limits defined by the states of Piauí, Goiás and Tocantins, and by the Corrente river basin and some basins of the middle reach of the São Francisco River. Its geographic coordinates are 10º10'S to 13º20'S and 43º08'W to 46º37'W. The drainage area of the catchment is about 75.000 km² and its topography is mainly characterized by abrupt altimetric amplitude changes in some sites.

Three types of climates can be found in the catchment: humid, in the western portion; sub-humid, in the central portion of the basin; and semiarid, in the east (Figure 1). The average annual temperature is 24.3 ºC. The precipitation regime is marked by strong seasonality, with the wet months between November and May. Mean annual rainfall ranges from 1998 mm, nearby the Ondas and the Fêmeas rivers headwaters, to 729 mm, in the
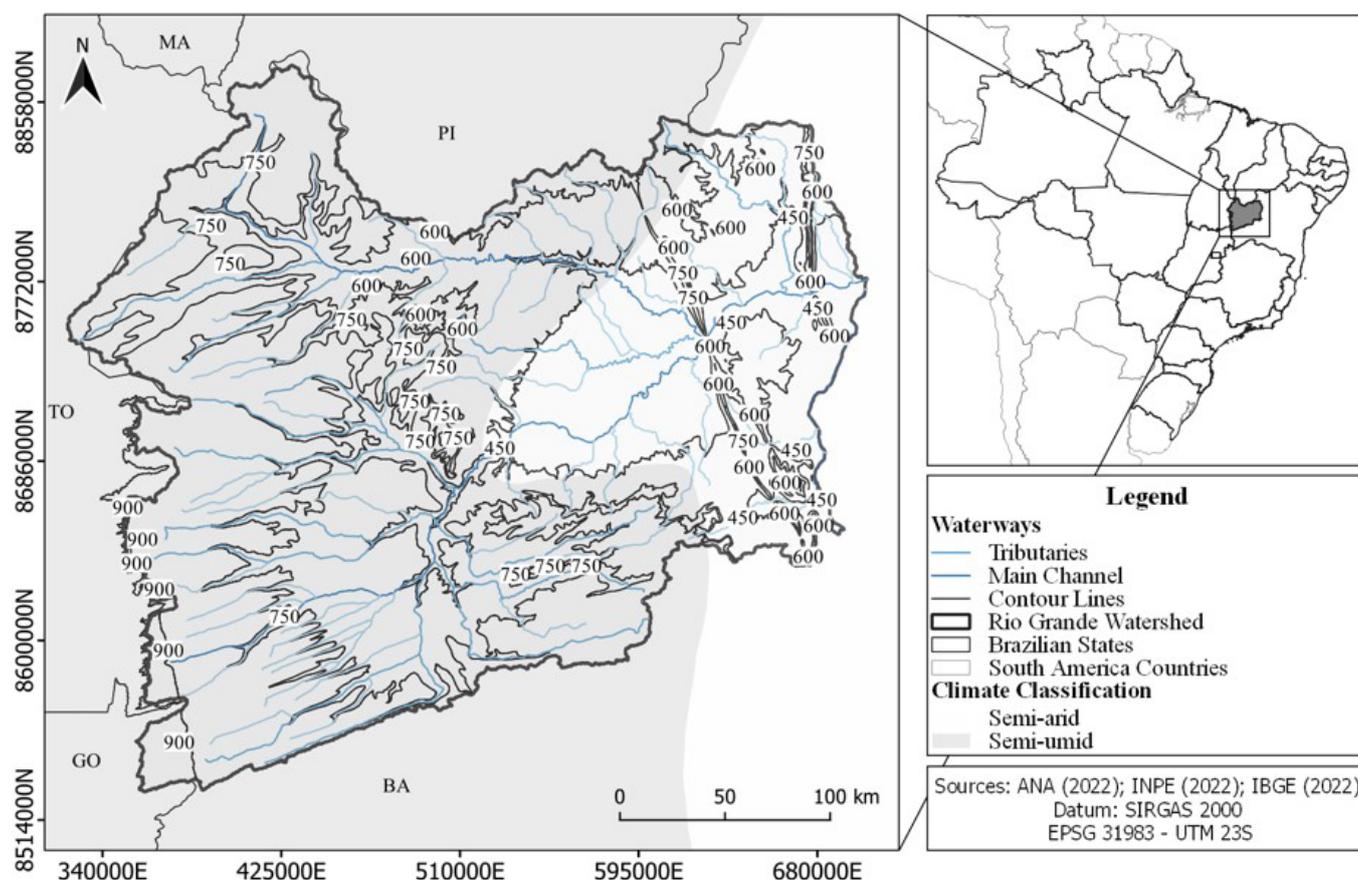
Figure 1. General aspects of Rio Grande watershed.

city of Barra-BA, at the confluence with the São Francisco River (Moreira & Silva, 2010).

Daily rainfall series from gauging stations in the region and 120 km away from its border, with at least 30 years of observations, as suggested by Naghettini (2017) for at site frequency analysis, were selected in the information system HidroWeb from the National Water Agency (Agência Nacional de Águas e Saneamento Básico, 2022). A simplified data quality check was performed to discard gauges with more than 20% of missing data (23 gauges) and/or time series with less than 30 years of records (86 gauges), as well as to eliminate unreasonably large daily rainfall amounts, defined by the sample 3^0 Quartile plus 1.5 times the empirical interquartile range, which could possibly comprise errors. We have also applied the Mann-Kendall non-parametric test, at the 5% significance level, to the reduced series of annual maximum rainfall in each station, for assessing the existence of monotonic trends; in 21 gauges, significant trends were indicated by the test, which has led to their exclusion. As a result, from the 175 rainfall gauging stations located in the region, only 45 were selected for subsequent analyses based on the outlined criteria (Figure 2).

Besides obtaining the timeseries of annual maximum rainfall events, it is necessary to determine the frequency associated with the intense precipitations by using probability distributions (Naghettini, 2017). Based on Extreme Values Theory, as discussed in Papalexiou and Koutsoyiannis (2013), rainfall block-maxima can lead to three types of domains of attraction, namey, the type-I or

Gumbel (light upper tail), the type-II or Fréchet (heavy upper tail), and the type-III or reversed Weibull (bounded from above). The Generalized Extreme Value distribution unifies all these domains into a single mathematical form.

Previous research has indicated that maximum rainfall amounts, at the daily time scale, are likely to be heavy-tailed variates (Opitz et al., 2018; Papalexiou & Koutsoyiannis, 2013; Serinaldi & Kilsby, 2014). However, when fitting the GEV distribution with the method of L-moments (Hosking & Wallis, 1997; Naghettini, 2017), our results seemed to converge to upper-bounded models, which, according to Costa and Fernandes (2017), might increase bias and uncertainty of the analysis when the upper bounds are defined without a physical basis. As a result, in this study, the Gumbel probability distribution was utilized for modeling the rainfall block-maxima in each gauging station. Parameter estimation was also performed with the method of L-moments (Hosking & Wallis, 1997; Naghettini, 2017). We note that, despite being a light-tailed model, the Gumbel distribution has been widely utilized in heavy rainfall studies (Aragão et al., 2013; Guimarães & Naghettini, 1998; Libertino et al., 2018; Manke et al., 2022; Moreira et al., 2020).

To evaluate the goodness-of-fit of the Gumbel distribution to the samples, the Anderson-Darling test was performed. We also computed the coefficient of determination (r^2) and the percent bias (p-bias) (Ferreira et al., 2020; Moriasi et al., 2007), given by Equation 1 and Equation 2, respectively, for objectively assessing the goodness-of-fit of the models.
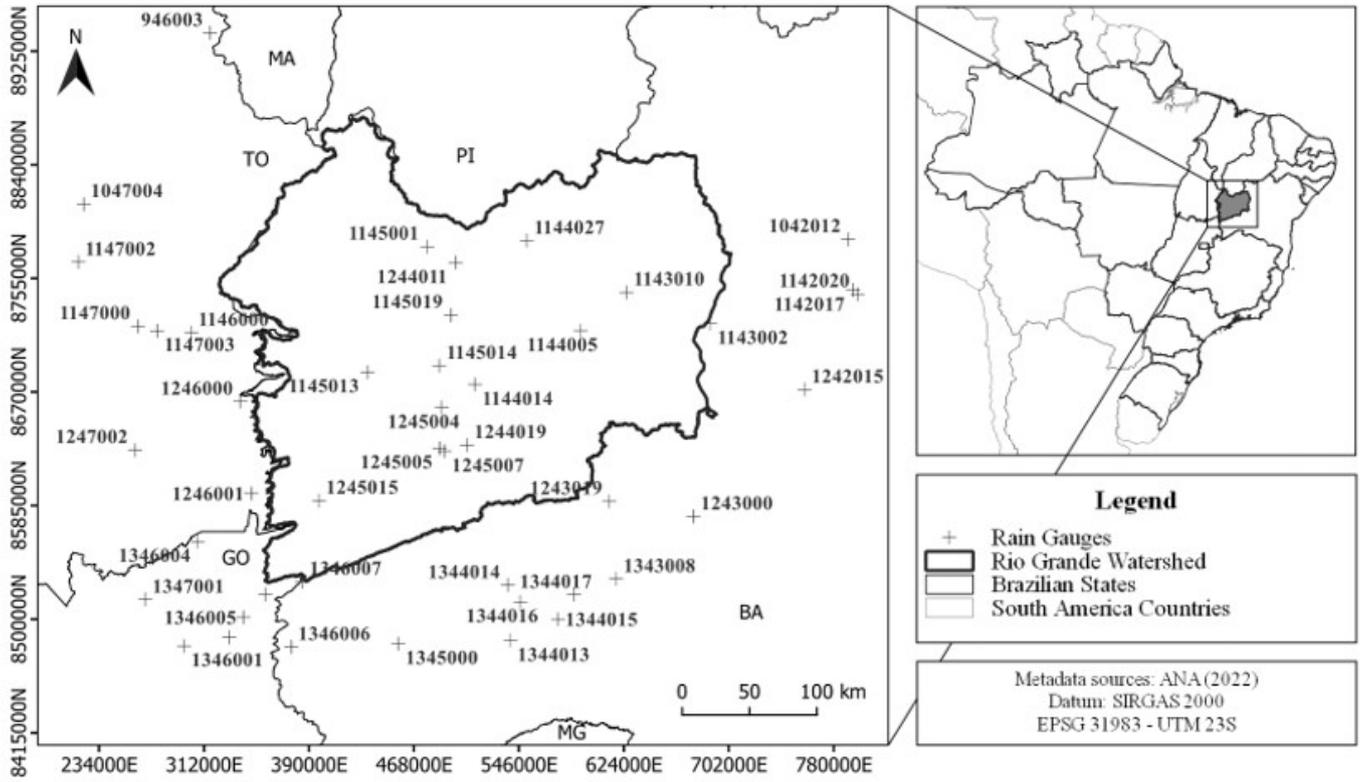
Figure 2. Grande river basin and selected rain gauges.

$$r^2 = \left[\frac{\sum_{i=1}^{N}\left(I_{obs(i)} - \overline{I_{obs}}\right) \cdot \left(I_{calc(i)} - \overline{I_{calc}}\right)}{\sqrt{\sum_{i=1}^{N}\left(I_{obs(i)} - \overline{I_{obs}}\right)^2 \cdot \sum_{i=1}^{N}\left(I_{calc(i)} - \overline{I_{calc}}\right)^2}} \right]^2 \quad (1)$$

$$p-bias = \frac{\sum_{i=1}^{N}\left(I_{cal(i)} - I_{obs(i)}\right)}{\sum_{i=1}^{N}\left(I_{obs(i)}\right)} \quad (2)$$

in which $I_{obs(i)}$ is the observed rainfall amounts in time period i, $I_{calc(i)}$ denotes the rainfall amounts obtained from the fitted model, $\overline{I_{obs}}$ is the mean of observed rainfall amount, $\overline{I_{calc}}$ is the mean of calculated intensities of precipitation and N is the total number of observations.

According to Koutsoyiannis et al. (1998), although not relying on a rigorous theoretical underpin for prescribing a suitable functional form, the IDF relationship may be characterized by quotient of power functions, in which the numerator depends solely on the return period T_r and the denominator depends on the duration or time scale d. Such a model is given by Equation 3.

$$I = \frac{\lambda \cdot Tr^{\kappa}}{(d + \theta)^{\eta}} \quad (3)$$

in which I is rainfall intensity (mm/h), Tr is return period (years), d is rainfall duration (hours) and λ, , θ and η are the model parameters.

For estimating the parameters of the IDF relationships, the block-maxima series, for the 1-day duration, were extracted in each site. Next, the quantiles associated with the return periods of 2, 5, 10, 20, 25, 50 and 100 years were estimated from the Gumbel quantile function. Then, the daily rainfall amounts, for each return period, were desegregated with the method suggested by the Departamento de Águas e Energia Elétrica (1980), which, despite lacking a rigorous mathematical background, has been extensively utilized in many similar studies (Aragão et al., 2013; Moreira et al., 2020; Souza et al., 2019), with acceptable errors with respect to the intensities derived from-pluviographic data in many regions in Brazil (Abreu et al., 2022; Dorneles et al., 2019). The disaggregation coefficients are presented in Table 1.

The parameters of the IDF model, λ, κ, θ and η were estimated using the Differential Evolution Algorithm for global optimization, with the package "*DEoptmin*" (Mullen et al., 2011) implemented in R language (R Core Team, 2021). The objective function comprised the minimization of the root mean square error (RMSE), shown in Equation 4.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(I_{obs(i)} - I_{calc(i)}\right)^2}{N - 1}} \quad (4)$$

Once the IDF parameters over the Grande River basin were obtained, they were spatialized using the IDW method. According to Pellicone et al. (2018), the variable of interest z (in this case, the IDF parameters) can be calculated at an ungauged site with Equation 5.

$$\hat{z} = \sum_{i=1}^{N} \left(\frac{d_i^{-p}}{\sum_{i=1}^{N} d_i^{-p}} \right) \cdot z_i \qquad (5)$$

in which z_i denotes the known value of variable z at a site i, N denotes the number of neighbor stations, d_i is the Euclidean distance from site i to ungauged site, and p is a control parameter. Following the recommendations of Pellicone et al. (2018) and Das and Wahiduzzaman (2022), we have adopted a value of 2 for parameter p, which entails a larger weight to nearby sites in the interpolation procedure.

For the spatialization of the quantiles, both the IDW method and ordinary kriging were considered, as suggested by Das and Wahiduzzaman (2022), for comparison. The ordinary kriging, which is based on a continuous Gaussian field with covariance function materialized by semivariograms, is the most common kriging method (Carvalho & Vieira, 2001; Libertino et al., 2018). For modeling purposes, the empirical semivariance, which is a discrete function used to measure the variability between pairs of points and is formally defined in Equation 6, and a theoretical semi-variograms must be computed (Das & Wahiduzzaman, 2022; Pellicone et al., 2018).

$$\hat{\gamma}(h) = \frac{1}{2 \cdot N(h)} \cdot \sum_{i=1}^{N(h)} \left(I(u_i) - I(u_i + h) \right)^2 \qquad (6)$$

in which h is the distance between two gauged sites, $\hat{\gamma}(h)$ is the semivariance, $N(h)$ is the number of pairs of sites at distance h, and $I(u_i)$ is the value of parameter in site u_i.

A theoretical semi-variogram, such as spherical and exponential models, is defined by the lag-distance in which the semi-variance is computed, the sill, the range and the nugget (Libertino et al., 2018). In this study, we have fitted the Spherical, the Gaussian and the Exponential semivariograms with the maximum likelihood method. Model selection was based on the minimum value of the Sum of Squared Errors (SSE), expressed by Equation 7.

$$SSE = \sum_{i=1}^{N} \left(I_{obs}(i) - I_{calc}(i) \right)^2 \qquad (7)$$

in which N is the number of sites.

For assessing the effectiveness of the spatialization procedure, we have resorted to cross-validation, as recommended by Pellicone et al. (2018) and Das and Wahiduzzaman (2022).

Table 1. Disaggregation coefficients to different rainfall durations.

Ratio of durations	Coefficient	Ratio of durations	Coefficient
24h / 1day	1.14	30min / 1h	0.74
12h / 1day	0.85	25min / 30 min	0.91
10h / 24h	0.82	20min / 30min	0.81
8h / 24h	0.78	15min / 30min	0.70
6h / 24h	0.72	10min / 30min	0.54
1h / 24h	0.42	5min / 30min	0.34

Source: Departamento de Águas e Energia Elétrica (1980).

Goodness-of-fit measures encompasses the RMSE and p-bias, which were computed with the empirical and the predicted rainfall intensities for events with duration of 24h and return period of 2 and 50 years. Finally, the quantiles associated with return periods of 2, 10, 25 and 50 years with a duration of 15 minutes and 12 hours were spatialized using the best model to provide examples of the overall methodology.

RESULTS AND DISCUSSION

We first fitted the Gumbel model to the reduced time-series of all 45 rainfall gauging stations. The goodness-of-fit was formally assessed by the Anderson-Darling test, which has high power in the tails of the distributions, and by the computation of the coefficient of determination (r^2) and the percent bias (p-bias).

The results in Figure 3a and 3b indicate that, for most gauging stations, the Gumbel model was able to capture the patterns of variability of the maximum daily rainfall amounts, although a slight tendency of systematic underestimation of the empirical quantiles was perceived from the value of p-bias. In addition, all the p-values resulting from Anderson-Darling test were larger than the significance level of the test, which also provides evidence of suitability of the Gumbel distribution for modeling the rainfall amounts in our study area.

After fitting the Gumbel probability distribution (parameters depicted in Figure 4), the desegregation method recommended by Departamento de Águas e Energia Elétrica (1980) was applied for providing estimates of the maximum rainfall amounts at sub daily time scales, considering the return periods of 2, 5, 10, 20, 25, 50 and 100 years. Then, the parameters of each IDF equation were estimated with the Differential Evolution Algorithm for global optimization. Results are summarized in Table 2, along with the goodness-of-fit assessment of the regression models.

As one may observe from Table 3 (and also in Figure 5), the parameters in the numerator of the IDF model, λ and κ, which control the decay of the upper tail of the Gumbel distribution,
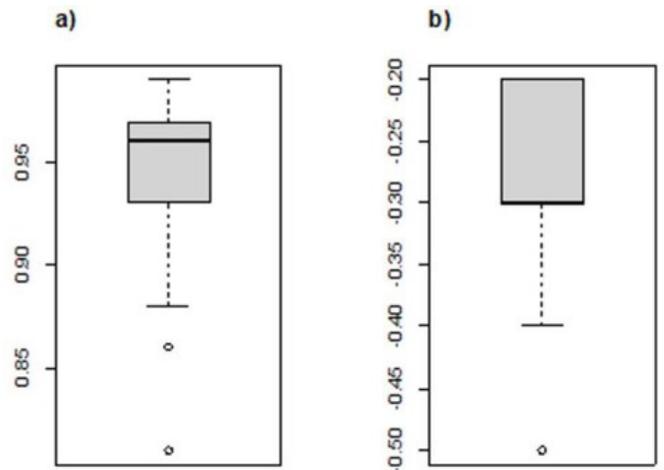
Figure 3. (a) coefficient of determination (r²) and (b) p-bias (%) for the goodness of fit between sample dada and quantiles of Gumbel probability distribution, for estimative of empirical plotting positions it was used Cunnane approximation.

Table 2. IDF parameters and goodness-of-fit metrics for the model in Equation 3.

Rain gauge	λ	κ	θ	η	RMSE	p-bias (%)	Coefficient of determination (r^2)
946003	38.3407	0.1580	0.1971	0.7580	2.48	0.00	0.9979
1042012	31.5480	0.1884	0.1971	0.7580	8.38	-9.80	0.9963
1047004	35.0560	0.2010	0.1971	0.7580	6.04	4.10	0.9945
1142017	29.6207	0.1641	0.1971	0.7580	17.91	-21.30	0.9979
1142020	30.9594	0.1656	0.1971	0.7580	14.63	-17.40	0.9979
1143002	36.8106	0.2028	0.1971	0.7580	10.27	9.90	0.9942
1143010	36.1907	0.1790	0.1971	0.7580	3.02	0.50	0.9972
1144005	37.2851	0.1727	0.1971	0.7580	3.14	1.60	0.9976
1144014	30.8785	0.2098	0.1971	0.7580	6.08	-5.80	0.9928
1144027	37.9695	0.1890	0.1971	0.7580	8.69	8.70	0.9962
1145001	33.3950	0.1761	0.1971	0.7580	6.95	-8.00	0.9974
1145013	33.9834	0.1547	0.1971	0.7580	10.56	-12.20	0.9978
1145014	36.2693	0.1503	0.1971	0.7580	6.93	-7.60	0.9976
1145019	35.5300	0.1735	0.1971	0.7580	3.42	-2.90	0.9976
1146000	35.4630	0.1878	0.1971	0.7580	3.70	1.20	0.9964
1147000	37.3212	0.1493	0.1971	0.7580	5.16	-5.20	0.9975
1147002	36.9048	0.1761	0.1971	0.7580	3.27	1.60	0.9974
1147003	37.3500	0.1459	0.1971	0.7580	5.91	-6.10	0.9973
1242015	27.5183	0.1830	0.1971	0.7580	18.80	-22.60	0.9969
1243000	32.2353	0.1785	0.1971	0.7580	8.95	-10.60	0.9973
1243019	37.3445	0.1768	0.1971	0.7580	4.07	3.00	0.9974
1244011	35.9524	0.1736	0.1971	0.7580	2.91	-1.70	0.9976
1244019	39.7183	0.1888	0.1971	0.7580	12.63	13.70	0.9962
1245004	35.3445	0.1736	0.1971	0.7580	3.68	-3.40	0.9976
1245005	38.5224	0.1415	0.1971	0.7580	4.91	-4.40	0.9968
1245007	37.5838	0.1421	0.1971	0.7580	6.39	-6.50	0.9969
1245015	35.5498	0.1862	0.1971	0.7580	3.53	0.90	0.9965
1246000	37.7133	0.1844	0.1971	0.7580	6.80	6.50	0.9967
1246001	36.0317	0.1714	0.1971	0.7580	3.02	-2.20	0.9977
1247002	32.8359	0.1723	0.1971	0.7580	9.01	-10.60	0.9977
1343008	33.6851	0.1733	0.1971	0.7580	6.94	-8.00	0.9976
1344013	35.8251	0.1757	0.1971	0.7580	2.88	-1.50	0.9975
1344014	34.6589	0.1904	0.1971	0.7580	3.50	-0.30	0.9960
1344015	37.9329	0.1872	0.1971	0.7580	8.07	8.00	0.9964
1344016	35.4643	0.1400	0.1971	0.7580	10.99	-12.40	0.9967
1344017	35.4002	0.1808	0.1971	0.7580	3.01	-1.10	0.9971
1345000	33.2569	0.2411	0.1971	0.7580	13.78	11.70	0.9851
1346000	31.8437	0.1492	0.1971	0.7580	16.31	-19.10	0.9975
1346001	34.9734	0.1673	0.1971	0.7580	5.60	-6.20	0.9979
1346002	34.7456	0.1871	0.1971	0.7580	3.31	-1.10	0.9964
1346004	37.7059	0.1491	0.1971	0.7580	4.53	-4.30	0.9975
1346005	32.1770	0.1908	0.1971	0.7580	6.55	-7.40	0.9960
1346006	35.7780	0.1556	0.1971	0.7580	6.63	-7.40	0.9978
1346007	36.1201	0.1852	0.1971	0.7580	4.02	2.20	0.9967
1347001	32.4031	0.1415	0.1971	0.7580	16.81	-19.60	0.9968

smoothly vary across space, with a few abrupt variations at specific locations. This is probably related to distinct climate and topographic conditions along the study region (Aragão et al., 2013; Campos et al., 2014, 2017), which might intensify storm bursts at short time scales. In contrast, the parameters in the denominator, η and θ, are constant for all rainfall gauging stations, which may be ascribed to the disaggregation method – in effect, all gauges are located in the same isozone defined by the Departamento de Águas e Energia Elétrica (1980). This might constitute a major limitation

of our approach because previous research (Blanchet et al., 2016; Innocenti et al., 2017) has indicated that the maximum rainfall scaling regime across durations is strongly affected by climate and terrain complexity and, hence, such parameters are expected to vary across large geographic regions.

Figure 6 depicts the goodness-of-fit assessment of the IDF models. The median values of RMSE, p-bias and r^2 are 6.07, -3.40 and 0.997, respectively, which, along with the relatively narrow interquartile amplitudes, might suggest an overall good agreement

between model estimates and the empirical evidence. We note, however, that, despite the very smooth spatial variation of the coefficient of determination, the values of RMSE and p-bias might be large at some gauging stations, which could indicate that
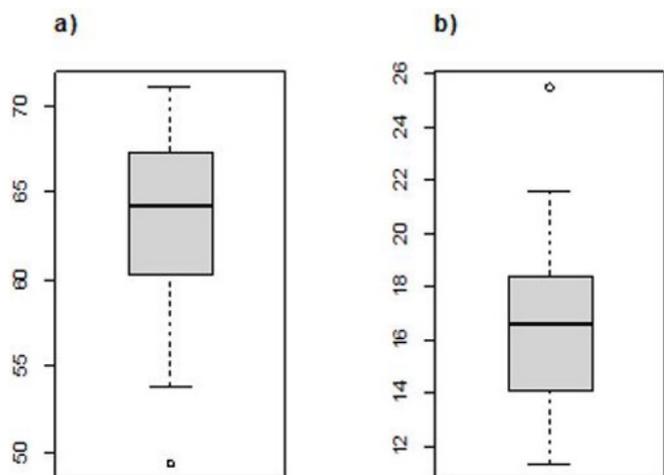
the modeled quantile curves have offsets with respect to the 1:1 line in these locations. Moreover, we could not associate these large deviations with climate or topographic gradients, which, at least to some extent, may suggest some lack of fit of the Gumbel model to the sample points in these gauges.

Figure 7 and Figure 8 illustrate the spatial variation of parameters λ and κ across the Rio Grande River catchment, as estimated with the IDW technique. Spatial patterns cannot be readily perceived for λ, which suggests that local features, such as more complex topography, might exert a stronger influence in the variation of the referred parameter than larger scale features such as climate conditions. For κ, however, a noticeable gradient develops in the southwestern-northeastern direction, with a possible exception to a relatively small region at the southern portion of the catchment. This gradient may suggest that climate is the main driver of the variation of κ, given that the western portion of the catchment is a semi-humid climate area whereas the eastern portion comprises a semi-arid region (Bahia, 2003).

The spatial patterns of variation of parameters λ and κ could not be properly captured with the ordinary kriging technique. In fact, irrespective of the semivariogram utilized for modeling the Gaussian field, such an approach resulted in oversmoothed surfaces, which could not represent local variations in the values of the parameters. The lack of fit might be, to some extent, attenuated by introducing covariates such as topography to the

Figure 4. (a) Gumbel location parameter and (b) Gumbel scale parameter obtained using L-moments [[Q1: Q1]].

Figure 5. Illustration of variation of IDF parameters (a) lambda and (b) kappa.

Table 3. Summary of results for ordinary kriging theoretical model selection.

Theoretical model	2-year return period			
	SSE	Nugget	Partial sill	Range
Spherical	$5.44 \cdot 10^{-12}$	0.00	0.0390	58984.2
Exponential	$6.91 \cdot 10^{-12}$	0.00	0.0428	33410.0
Gaussian	$3.58 \cdot 10^{-12}$	0.00	0.0396	25808.6
Theoretical model	50-year return period			
	SSE	Nugget	Partial sill	Range
Spherical	$3.33 \cdot 10^{-10}$	0.00	0.2676	31545.8
Exponential	$4.33 \cdot 10^{-10}$	0.00	0.2691	13125.0
Gaussian	$1.27 \cdot 10^{-10}$	0.21	0.3928	9289.8

Figure 6. Boxplot of goodness-of-fit metrics of (a) RMSE; (b) p-bias (%) and (c) coefficient of determination between Gumbel quantiles and IDF estimations.
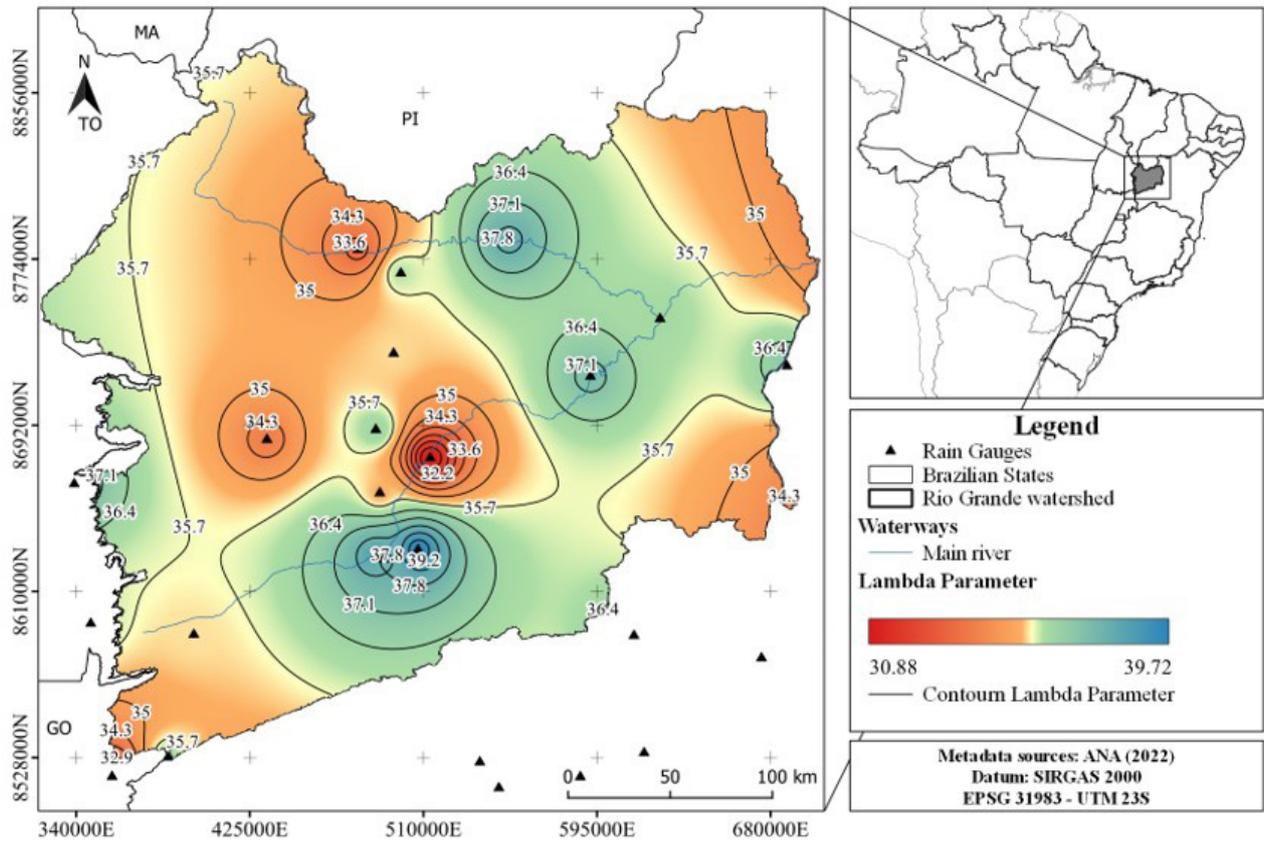
Figure 7. Spatialization of lambda IDF parameter for Rio Grande watershed.
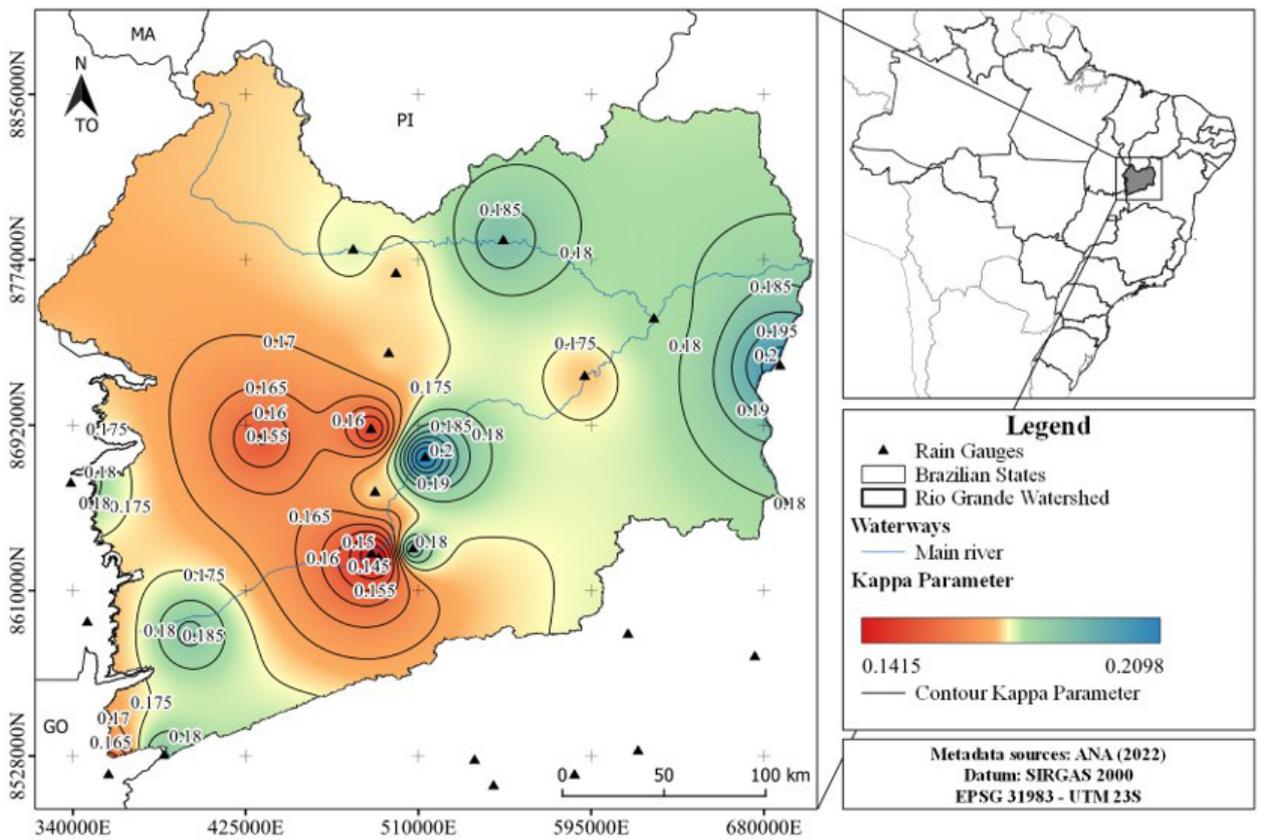

Figure 8. Spatialization of kappa IDF parameter for Rio Grande watershed.

kriging procedure (regression-kriging), but this assessment is beyond the scope of this study and is left for future work.

We then proceed to the spatialization of maximum rainfall quantiles, according to the procedures discussed in the previous section – apparently, the spatial variation on the quantiles is much smoother than the IDF parameters, which enables the use of ordinary kriging for obtaining models with suitable predictive skills. A summary of the parameters of the semivariograms fitted with the kriging technique, for the return periods of 2 and 50 years, is provided in Table 3.

Overall, the values of SSE are not much different, but the Gaussian semivariogram consistently outperformed the Exponential and the Spherical counterparts and it was hence selected for subsequent analyses. We note that the Gaussian semivariogram presented the smaller values for the range in both situations, which strongly reduces the influence of more distant sites in prediction. Also, the Gaussian model was the only one to exhibit a nugget effect for the 50-year return period. This could be ascribed to the more pronounced distinctions in the estimates belonging in the upper tail of the Gumbel distribution, which may then manifest themselves as "local effects" during the interpolation.

Table 4 presents the results of quantile spatialization, for both IDW and the ordinary kriging techniques, under calibration and cross-validation. For both the 2-year and 50-year return periods, the values of RMSE and p-bias indicate that the IDW method provided slightly closer quantile estimates to the empirical information than the ordinary kriging, although both approaches are, to a very small extent, biased towards underestimation. Since the IDW model is less complex and performs very similarly to the ordinary kriging approach, we selected it for quantile spatialization.

The IDW interpolations, for return periods of 2, 10, 25, 50 years and duration of 30 min and 12 hours, are presented in Figure 9 and in Figure 10, respectively. By analyzing Figure 9 and Figure 10, one may notice that the spatial distribution of the quantiles, for all return periods, resembles that of climate classification (see Figure 1), with a strong gradients of maximum

Table 4. Summary of cross-validation results comparing both IDW and Ordinary Kriging methods.

Return period	Step	IDW		Ordinary Kriging	
		RMSE	p-bias	RMSE	p-bias
2 - years	Calibration	0.0001	0.00%	0.0017	0.01%
	Cross-validation	0.1915	-0.29%	0.2348	-0.79%
50 - years	Calibration	0.0003	0.00%	0.1577	-0.01%
	Cross-validation	0.4699	-1.62%	0.4788	-0.54%
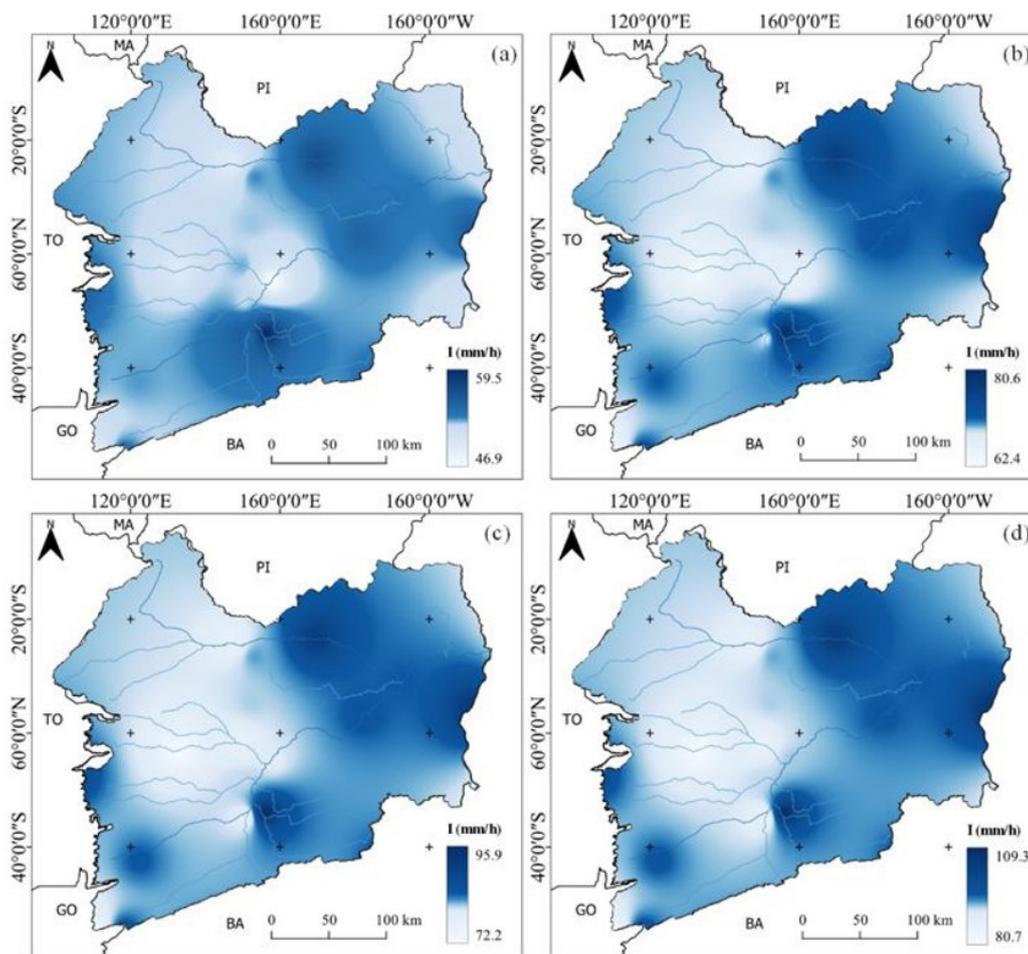
Figure 9. Spatialization of quantiles associated a precipitation with duration of 30 minutes and return periods of (a) 2 years; (b) 10 years; (c) 25 years and (d) 50 years.
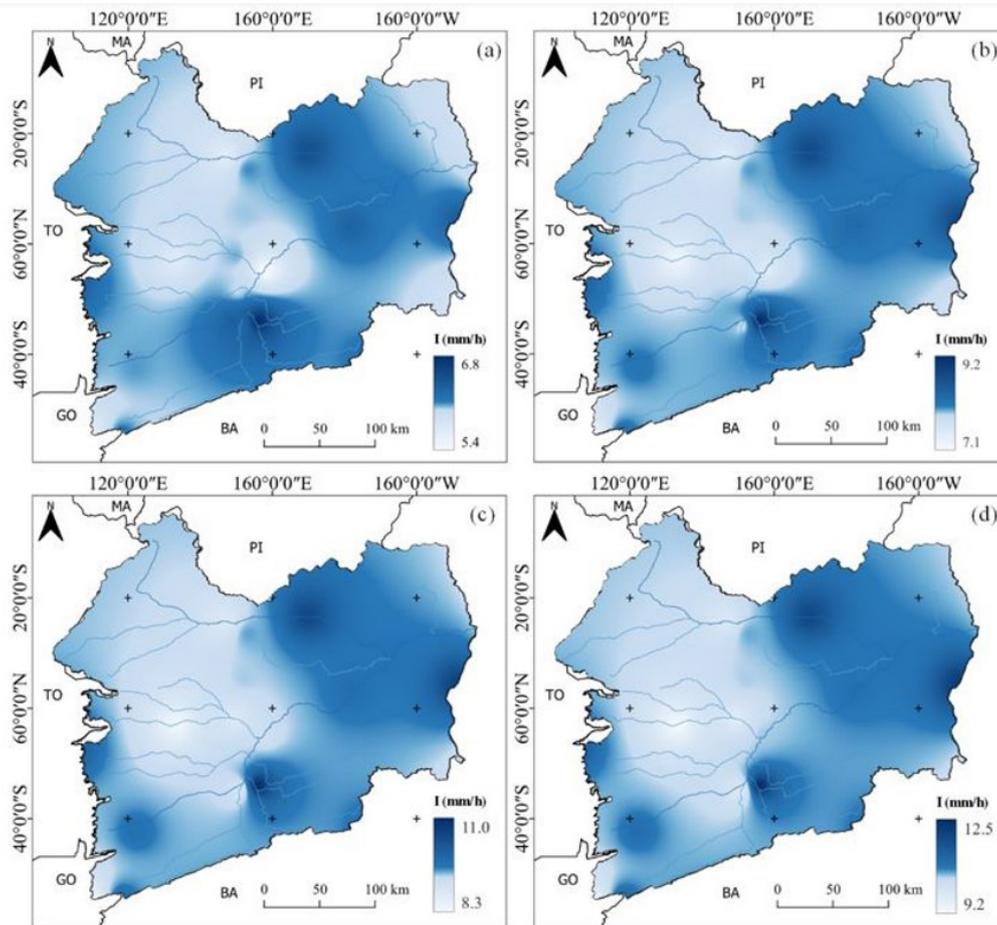
Figure 10. Spatialization of quantiles associated a precipitation with duration of 12 hours and return periods of (a) 2 years; (b) 10 years; (c) 25 years and (d) 50 years.

rainfall intensity in the southwestern-northeastern direction. For a given exceedance probability, the more intense events occur in the semi-arid region, but the gradients do not considerably vary with the return period. More abrupt variations are perceived for the short duration rainfall amounts (Figure 9), which may be related to local orographic influence and are probably a result of the rougher surface obtained for parameter λ of the IDF equation (which was herein utilized for quantile estimation). As the durations increase (Figure 10), smoother surfaces are obtained. We should stress, however, that the possibly distinct scaling laws across durations in large regions, which could affect the spatial distribution of the maximum rainfall intensities (e.g., Ghanmi et al., 2016), are not taken into account by our IDF model. As a result, some level of bias is expected in the estimation of the spatial surfaces, which should demand careful inspections by practitioners when designing hydraulic structures on the basis of our regional model.

CONCLUSIONS

This paper discussed the estimation of regional Intensity-Duration-Frequency (IDF) relationships in the Grande River Catchment, a relatively poorly gauged area which is located in the Brazilian state of Bahia. For this, maximum daily rainfall amounts were retrieved from 45 gauges inside and in the vicinity of the catchment and disaggregated to sub daily times scales according to the procedure suggested by the Departamento de Águas e Energia Elétrica (1980). Then, spatial surfaces for the parameters of the IDF models and for some quantiles of interest for design purposes were obtained by the Inverse Distance Weighting (IDW) and the ordinary kriging techniques.

Our results suggest that, as opposed to some indications in the literature (e.g., Koutsoyiannis, 2021; Papalexiou & Koutsoyiannis, 2013) the maximum rainfall amounts in the study region can be reasonably described by a light-tailed distribution, namely, the Gumbel distribution, at the daily time scale (in fact, empirical evidence suggested upper-bounded variates for most gauges and heavy tailed ones for a few gauges). As a matter of fact, the use of the Gumbel distribution as a regional model has led to low levels of bias (<10% in absolute value) and did not entail large errors for high return levels. A potential advantage of utilizing the Gumbel model is that, as no shape parameter needs to be inferred from finite and usually small samples (as in the GEV case), the estimation errors in the tail index should not propagate during the disaggregation process, which could lead to unreasonably large intensities for short duration storms. Nonetheless, as the Gumbel model does not comprise a power law, the scaling properties across durations may not hold (Koutsoyiannis, 2021), and this could certainly affect the disaggregated rainfall amounts.

The spatialization processes indicated that, for our dataset, the IDW technique was able to model both the IDF parameters and rainfall quantiles. Ordinary kriging, in turn, could not represent the spatial variation of the IDF parameters, but performed as well as the IDW approach for the rainfall quantiles – the latter originated smoother spatial surfaces. This limitation might be related to the questionable assumption of a common mean for the Gaussian field in an area with very distinct climate conditions and terrain complexity. Also, the abrupt variations of parameter λ might not be described by the smooth surface that stemmed from the semivariograms. We hypothesize that including covariates such as topography might improve the prediction abilities of the geostatistical model, and this framework will be contemplated in our next research developments.

Finally, some remarks should be made regarding our simplified approach for rainfall disaggregation. In fact, despite the simplicity and the widespread use, the disaggregation approach by the Departamento de Águas e Energia Elétrica (1980) does not properly capture the spatial variation in the scaling regime of the maximum rainfall intensities – from a physical point of view, it is not reasonable to assume that the same scale exponent and duration offset are valid over such a large region (Innocenti et al., 2017). This fact might affect quantile estimation and risk assessment. However, disaggregation of daily rainfall amounts for finer times scales is still an active research topic (see the discussions in Diez-Sierra and del Jesus (2019) and Aguilar and Costa (2020)), and our approach might be improved in future work if high-resolution rainfall data is made available. Still, practitioners may benefit from our results for indirectly estimating floods and designing the correspondent conveyance structures at ungauged sites in the Grande River catchment.

ACKNOWLEDGEMENTS

REFERENCES

Abreu, M. C., Almeida, L. T., Silva, F. B., Fraga, M. S., Lyra, G. B., Souza, A., Silva, D. D., & Cecílio, R. A. (2022). Disaggregation coefficients for obtaining rainfall intensity-duration-frequency curves: concepts, models, errors and trends in Minas Gerais, Brazil. *Urban Water Journal*, 1-14. http://dx.doi.org/10.1080/1573062X.2022.2105237.

Agência Nacional de Águas e Saneamento Básico – ANA. (2022, abril 18). *HIDROWEB. Séries Históricas de Estações*. Retrieved in 2022, September 3, from https://www.snirh.gov.br/hidroweb/serieshistoricas

Aguilar, M. G., & Costa, V. A. F. (2020). A regional similarity-based approach for sub-daily rainfall nonparametric generation. *Revista Brasileira de Recursos Hídricos*, *25*, e5. https://doi.org/10.1590/2318-0331.252020190054.

Aragão, R., Santana, G. R., Costa, C. E. F. F., Cruz, M. A. S., Figueiredo, E. E., & Srinivasan, V. S. (2013). Chuvas intensas para o estado de Sergipe com base em dados desagregados de chuva diária. *Revista Brasileira de Engenharia Agrícola e Ambiental*, *17*(3), 243-252. http://dx.doi.org/10.1590/S1415-43662013000300001.

Asadieh, B., & Krakauer, N. Y. (2017). Global change in streamflow extremes under climate change over the 21st century. *Hydrology and Earth System Sciences*, *21*(11), 5863-5874. http://dx.doi.org/10.5194/hess-21-5863-2017.

Back, Á. J. (2020). Alternative model of intense rainfall equation obtained from daily rainfall disaggregation. *RBRH*, *25*, e2. http://dx.doi.org/10.1590/2318-0331.252020190031.

Bahia. (2003). *Diagnóstico e Regionalização (Plano Estadual de Recursos Hídricos do Estado da Bahia)*. Salvador: SEINFRA-BA e SRH-BA.

Blanchet, J., Ceresetti, D., Molinié, G., & Creutin, J.-D. (2016). A regional GEV scale-invariant framework for Intensity–Duration–Frequency analysis. *Journal of Hydrology*, *540*, 82-95. http://dx.doi.org/10.1016/j.jhydrol.2016.06.007.

Bougara, H., Hamed, K. B., Borgemeister, C., Tischbein, B., & Kumar, N. (2020). Analyzing trend and variability of rainfall in the Tafna Basin (Northwestern Algeria). *Atmosphere*, *11*(4), 347. http://dx.doi.org/10.3390/atmos11040347.

Campos, A. R., Santos, G. G., Silva, J. B. L., Irene Filho, J., & Loura, D. S. (2014). Equações de intensidade-duração-frequência de chuvas para o estado do Piauí. *Revista Ciência Agronômica*, *45*(3), 488-498. http://dx.doi.org/10.1590/S1806-66902014000300008.

Campos, A. R., Silva, J. B. L., Santos, G. G., Ratke, R. F., & Aquino, I. O. (2017). Estimate of intense rainfall equation parameters for rainfall stations of the Paraíba State, Brazil. *Pesquisa Agropecuária Tropical*, *47*(1), 15-21. http://dx.doi.org/10.1590/1983-40632016v4743821.

Carvalho, J. R. P., & Vieira, S. R. (2001). *Avaliação e comparação de estimadores-de-krigagem para variáveis agronômicas: uma proposta* (1ª ed, Vol. 13). Campinas: Embrapa Informática Agropecuária.

Costa, V., & Fernandes, W. (2017). Bayesian estimation of extreme flood quantiles using a rainfall-runoff model and a stochastic daily rainfall generator. *Journal of Hydrology*, *554*, 137-154. http://dx.doi.org/10.1016/j.jhydrol.2017.09.003.

Das, S., & Wahiduzzaman, M. (2022). Identifying meaningful covariates that can improve the interpolation of monsoon rainfall in a low-lying tropical region. *International Journal of Climatology*, *42*(3), 1500-1515. http://dx.doi.org/10.1002/joc.7316.

Departamento de Águas e Energia Elétrica – DAEE. Companhia de Tecnologia de Saneamento Ambiental – CETESB. (1980). *Drenagem urbana: manual de projeto* (2ª ed). São Paulo: DAEE & CETESB.

Diez-Sierra, J., & del Jesus, M. (2019). Subdaily rainfall estimation through daily rainfall downscaling using Random Forests in Spain. *Water*, *11*(1), 125. http://dx.doi.org/10.3390/w11010125.

Dorneles, V. R., Damé, R. C. F., Teixeira-Gandra, C. F. A., Méllo, L. B., Ramirez, M. A. A., & Manke, E. B. (2019). Intensity-Duration-Frequency relationships of rainfall through the technique of disaggregation of daily rainfall. *Revista Brasileira de Engenharia Agrícola e Ambiental*, *23*(7), 506-510. http://dx.doi.org/10.1590/1807-1929/agriambi.v23n7p506-510.

Ferreira, P. M. L., Paz, A. R., & Bravo, J. M. (2020). Objective functions used as performance metrics for hydrological models: state-of-the-art and critical analysis. *Revista Brasileira de Recursos Hídricos*, *25*, e42. https://doi.org/10.1590/2318-0331.252020190155.

Ghanmi, H., Bargaoui, Z., & Mallet, C. (2016). Estimation of intensity-duration-frequency relationships according to the property of scale invariance and regionalization analysis in a Mediterranean coastal area. *Journal of Hydrology*, *541*, 38-49. http://dx.doi.org/10.1016/j.jhydrol.2016.07.002.

Gonçalves, R. D., Stollberg, R., Weiss, H., & Chang, H. K. (2020). Using GRACE to quantify the depletion of terrestrial water storage in Northeastern Brazil: the Urucuia Aquifer System. *The Science of the Total Environment*, *705*, 135845. http://dx.doi.org/10.1016/j.scitotenv.2019.135845.

Guimarães, M., & Naghettini, M. (1998). Análise regional de frequência e distribuição temporal das tempestades na Região Metropolitana de Belo Horizonte - RMBH. *Revista Brasileira de Recursos Hídricos*, *3*(4), 73-88. http://dx.doi.org/10.21168/rbrh.v3n4.p73-88.

Hosking, J. R. M., & Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511529443.

Innocenti, S., Mailhot, A., & Frigon, A. (2017). Simple scaling of extreme precipitation in North America. *Hydrology and Earth System Sciences*, *21*(11), 5823-5846. http://dx.doi.org/10.5194/hess-21-5823-2017.

Jesus, J. B., & Nascimento, Y. S. (2020). Tempo de retorno e espacialização das precipitações máximas pelo método dos momentos para o estado da Bahia. *Engenharia Sanitaria e Ambiental*, *25*(1), 127-131. http://dx.doi.org/10.1590/s1413-41522020200508.

Koutsoyiannis, D. (2021). *Stochastics of hydroclimatic extremes: a cool look at risk* (1st ed). Athens: Kallipos.

Koutsoyiannis, D., Kozonis, D., & Manetas, A. (1998). A mathematical framework for studying rainfall intensity-duration-frequency relationships. *Journal of Hydrology*, *206*(1–2), 118-135. http://dx.doi.org/10.1016/S0022-1694(98)00097-3.

Libertino, A., Allamano, P., Laio, F., & Claps, P. (2018). Regional-scale analysis of extreme precipitation from short and fragmented records. *Advances in Water Resources*, *112*, 147-159. http://dx.doi.org/10.1016/j.advwatres.2017.12.015.

Ly, S., Charles, C., & Degré, A. (2013). Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: a review. *Biotechnologie, Agronomie, Société et Environnement*, *17*(2), 392-406.

Manke, E. B., Teixeira-Gandra, C. F. A., Damé, R. C. F., Nunes, A. B., Neta, M. C. C. C., & Karsburg, R. M. (2022). Seasonal Intensity-Duration-Frequency relationships for Pelotas, Rio Grande do Sul, Brazil. *Revista Brasileira de Engenharia Agrícola e Ambiental*, *26*(2), 85-90. http://dx.doi.org/10.1590/1807-1929/agriambi.v26n2p85-90.

Moreira, M. C., & Silva, D. D. (2010). *Atlas hidrológico da bacia hidrográfica do Rio Grande*. Porto Alegre: Gazeta Santa Cruz.

Moreira, P. G., Silva, J. B. L., Bento, N. L., Silva, D. P., & Souza, B. S. (2020). Estimativa dos parâmetros de equações Intensidade-Duração-Frequência de chuvas intensas para o Estado da Bahia, Brasil. *Revista Eletrônica do PRODEMA*, *14*(1), 14. http://dx.doi.org/1022411/rede2020.1401.10.

Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885-900. http://dx.doi.org/10.13031/2013.23153.

Mullen, K., Ardia, D., Gil, D., Windover, D., & Cline, J. (2011). DEoptim: an R Package for global optimization by differential evolution. *Journal of Statistical Software*, *40*(6), http://dx.doi.org/10.18637/jss.v040.i06.

Naghettini, M. (Org.). (2017). *Fundamentals of statistical hydrology*. Switzerland: Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-43561-9.

Nunes, A. A., Pinto, E. J. A., Baptista, M. B., Paula, M. H., & Xavier, M. O. (2021). Intensity-Duration-Frequency curves in the municipality of Belo Horizonte from the perspective of non-stationarity. *Revista Brasileira de Recursos Hídricos*, *26*, e29. https://doi.org/10.1590/2318-0331.262120210017.

Opitz, T., Huser, R., Bakka, H., & Rue, H. (2018). INLA goes extreme: Bayesian tail regression for the estimation of high spatio-temporal quantiles. *arXiv*, 1802.01085.http://dx.doi.org/10.48550/ARXIV.1802.01085.

Papalexiou, S. M., & Koutsoyiannis, D. (2013). Battle of extreme value distributions: a global survey on extreme daily rainfall: survey

on extreme daily rainfall. *Water Resources Research*, *49*(1), 187-201. http://dx.doi.org/10.1029/2012WR012557.

Pellicone, G., Caloiero, T., Modica, G., & Guagliardi, I. (2018). Application of several spatial interpolation techniques to monthly rainfall data in the Calabria region (southern Italy). *International Journal of Climatology*, *38*(9), 3651-3666. http://dx.doi.org/10.1002/joc.5525.

Pereira, S. B., Pruski, F. F., Silva, D. D., & Ramos, M. M. (2007). Estudo do comportamento hidrológico do Rio São Francisco e seus principais afluentes. *Revista Brasileira de Engenharia Agrícola e Ambiental*, *11*(6), 615-622. http://dx.doi.org/10.1590/S1415-43662007000600010.

Pizarro, R., Ingram, B., Gonzalez-Leiva, F., Valdés-Pineda, R., Sangüesa, C., Delgado, N., García-Chevesich, P., & Valdés, J. (2018). WEBSEIDF: a web-based system for the estimation of IDF curves in Central Chile. *Hydrology*, *5*(3), 40. http://dx.doi.org/10.3390/hydrology5030040.

R Core Team. (2021). R*: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved in 2022, September 3, from https://www.R-project.org/

Rabelo, A. E. C. G. C., Ribas, L. V. S., Coutinho, A. P., Ribeiro Neto, A., & Antonino, A. C. D. (2018). Espacialização dos parâmetros de equações de chuvas intensas para a Região Metropolitana do Recife. *Revista Brasileira de Geografia Física*, *11*(4), 1542-1554. http://dx.doi.org/10.26848/rbgf.v11.4.p1542-1554.

Sabino, M., Souza, A. P., Uliana, E. M., Lisboa, L., Almeida, F. T., & Zolin, C. A. (2020). Intensity-Duration-Frequency of maximum rainfall in Mato Grosso State. *Revista Ambiente & Água*, *15*(1), 1. http://dx.doi.org/10.4136/ambi-agua.2373.

Serinaldi, F. (2010). Multifractality, imperfect scaling and hydrological properties of rainfall time series simulated by continuous universal multifractal and discrete random cascade models. *Nonlinear Processes in Geophysics*, *17*(6), 697-714. http://dx.doi.org/10.5194/npg-17-697-2010.

Serinaldi, F., & Kilsby, C. G. (2014). Rainfall extremes: toward reconciliation after the battle of distributions. *Water Resources Research*, *50*(1), 336-352. http://dx.doi.org/10.1002/2013WR014211.

Souza, G. R., Bello, I. P., Oliveira, L. F. C., & Corrêa, F. V. (2019). Heavy rainfall maps in Brazil to 5-year return period. *Revista Ambiente & Água*, *14*(5), 1. http://dx.doi.org/10.4136/ambi-agua.2403.

Authors contributions

Alan de Gois Barbosa: Methodology definition, statistical procedures, spatialization procedures, results analysis, article writing.

Izaias Rodrigues de Souza Neto: Methodology definition, statistical procedures, article writing.

Veber Afonso Figueiredo Costa: Methodology definition, results analysis, article writing, supervision.

Ludmilson Abritta Mendes: Methodology definition, results analysis, article writing, supervision.

Editor-in-Chief: Adilson Pinheiro

Associated Editor: Carlos Henrique Ribeiro Lima