

UNIVERSIDADE FEDERAL DE ALFENAS - UNIFAL/MG

SANDRA VALÉRIA COELHO DA SILVA

ANÁLISE BAYESIANA DA PRECIPITAÇÃO MÁXIMA DE PETRÓPOLIS-RJ

Alfenas-MG
2021

SANDRA VALÉRIA COELHO DA SILVA

ANÁLISE BAYESIANA DA PRECIPITAÇÃO MÁXIMA DE PETRÓPOLIS-RJ

Dissertação apresentada ao Programa de Pós-Graduação em Estatística Aplicada e Biometria, área de concentração em Estatística Aplicada e Biometria da Universidade Federal de Alfenas (UNIFAL-MG), como parte dos requisitos para a obtenção do título de Mestre. Linha de Pesquisa: Modelagem Estatística e Estatística Computacional
Orientador: Prof. Dr. Luiz Alberto Beijo.
Coorientador: Prof. Dr. Joel Augusto Muniz.

Alfenas-MG
2021

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da Universidade Federal de Alfenas
Biblioteca Central – Campus Sede

S586a Silva, Sandra Valéria Coelho da
Análise bayesiana da precipitação máxima de Petrópolis-RJ / Sandra
Valéria Coelho da Silva– Alfenas, MG, 2021.
44 f.: il. –

Orientador: Luiz Alberto Beijo.
Dissertação (Mestrado em Estatística Aplicada e Biometria) – Universidade
Federal de Alfenas, 2021.
Bibliografia.

1. Chuva extrema. 2. Distribuição GEV. 3. Priori informativa. 4. Tempos de
retorno. I. Beijo, Luiz Alberto. II. Título.

CDD- 519.5

SANDRA VALÉRIA COELHO DA SILVA

ANÁLISE BAYESIANA DA PRECIPITAÇÃO MÁXIMA DE PETRÓPOLIS-RJ

A Banca examinadora abaixo-assinada aprova a Dissertação apresentada como parte dos requisitos para a obtenção do título de Mestre em Estatística Aplicada e Biometria pela Universidade Federal de Alfenas. Área de concentração: Estatística Aplicada e Biometria.

Aprovada em: 29 de julho de 2021.

Prof. Dr. Luiz Alberto Beijo
Instituição: Universidade Federal de Alfenas - UNIFAL-MG

Profa. Dra. Kelly Cristina Mota Gonçalves
Instituição: Universidade Federal do Rio de Janeiro - UFRJ

Prof. Dr. Fabrício Goeking Avelar
Instituição: Universidade Federal de Alfenas - UNIFAL-MG



Documento assinado eletronicamente por **Luiz Alberto Beijo, Professor do Magistério Superior**, em 29/07/2021, às 11:09, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Kelly Cristina Mota Gonçalves, Usuário Externo**, em 29/07/2021, às 11:13, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Fabrício Goeking Avelar, Professor do Magistério Superior**, em 29/07/2021, às 11:28, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.unifal-mg.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0564979** e o código CRC **9D00B850**.

AGRADECIMENTOS

À minha família, por todos os momentos de apoio, de descontração e de carinho.

A todos os meus amigos, por fazerem parte da minha vida e torná-la muito melhor.

Aos colegas das escolas em que trabalhei e trabalho: Colégio Estadual Monteiro Lobato e Ciep 465 Dr. Amílcar Pereira Da Silva, pela amizade e por ajudarem no meu processo de formação como profissional.

E aos amigos do IPEA, IBGE por me proporcionar a oportunidade de fazer parte em instituições tão renomadas.

A todos os professores que participaram da minha formação, desde o ensino fundamental até a pós - graduação.

À Universidade Federal de Alfenas, aos professores e servidores do programa de Pós- Graduação em Estatística Aplicada e Biometria, pela oportunidade concedida para a realização do mestrado. O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Ao meu orientador Luiz Alberto Beijo por todo o suporte dado durante todo o período desta dissertação.

A todas as pessoas que, de alguma forma, deram sua parcela de contribuição neste trabalho e de toda uma vida acadêmica.

RESUMO

A cidade de Petrópolis, situada na região serrana do estado do Rio de Janeiro-RJ, sofre frequentemente com estragos provocados por fortes chuvas, como as ocorridas nos anos de 2011 e 2013. Portanto, analisar e prever a ocorrência de precipitações pluviiais extremas em Petrópolis é fundamental para o planejamento de atividades vulneráveis a sua ocorrência, tais como a agropecuária e a remoção de pessoas das áreas de risco, sendo a primeira fundamental para a subsistência de muitas famílias locais e a segunda de extrema importância para evitar a perda de vidas humanas. A modelagem do comportamento desse evento extremo é feita, geralmente, por meio da distribuição Generalizada de Valores Extremos (GEV). A metodologia Bayesiana tem apresentado bons resultados na estimação dos parâmetros da distribuição GEV. Sendo assim, objetivou-se ajustar a distribuição GEV às séries históricas de precipitação máxima de Petrópolis, e avaliar diferentes estruturas de *prioris*, informativas e não informativas, na previsão dos valores de precipitação extrema de Petrópolis em diferentes tempos de retorno. Foi possível concluir que a distribuição a *priori* fundamentada em informações da cidade de Teresópolis, uma vez a variância, forneceu resultados mais precisos e acurados na predição dos níveis de retorno para a cidade de Petrópolis. Dessa forma, foram realizadas as predições para as precipitações máximas esperadas para os tempos de retorno de 2, 5, 10, 25, 50, 100, 115 e 135 anos utilizando-se essa estrutura de *priori*. Espera-se que, em um tempo médio de 5 anos, ocorra pelo menos um dia com precipitação máxima maior ou igual 100,7mm em Petrópolis-RJ.

Palavras-chave:

Chuva extrema; distribuição GEV; *Priori informativa*; tempos de retorno.

ABSTRACT

The city of Petrópolis, located in the mountainous region of the state of Rio de Janeiro (RJ), Brazil, frequently suffers from damage caused by heavy rains, such as those occurred in 2011 and 2013. Therefore, analyze and predict the occurrence of extreme rainfall in Petrópolis is fundamental for planning activities vulnerable to their occurrence, such as farming and the removal of people from risk areas, the first being fundamental for the livelihoods of many local families and the second being extremely important to avoid the loss of human life. The modeling of the behavior of this extreme event is usually done through the Generalized Distribution of Extreme Values (GEV). The Bayesian methodology has shown good results in estimating the parameters of the GEV distribution. Thus, the purpose was to adjust the GEV distribution to the historical series of maximum rainfall in Petrópolis, and to evaluate different structures of *prioris*, informative and non-informative in the forecasting of extreme rainfall values in Petrópolis at different return times. It was possible to conclude that the *priori* distribution based on information from the city of Teresópolis, given the variance, provided more accurate and accurate results in the prediction of return levels for the city of Petrópolis. In this way, forecasts for the maximum rainfall expected for the return times of 2, 5, 10, 25, 50, 100, 115 and 135 years were made using this *priori* structure. It is expected that, in an average time of 5 years, there will be at least one day with maximum rainfall greater than or equal to 100.7mm in Petrópolis.

Keywords: Extreme rain; GEV distribution; informative *Prior*; return times.

LISTA DE ILUSTRAÇÕES

Figura 1 – Mapa da Região Sudeste brasileiro.	27
Figura 2 – Diagrama de dispersão da precipitação máxima anual (1943 à 2020) das cidades de Petrópolis-RJ, Teresópolis-RJ e Juiz de Fora-MG. . .	31

LISTA DE TABELAS

Tabela 1 – Estimativas dos parâmetros da distribuição GEV via inferência Bayesiana e respectivas variâncias e covariâncias.	28
Tabela 2 – Estatísticas descritivas da precipitação máxima anual (mm) de Petrópolis-RJ, Teresópolis-RJ e Juiz de Fora-MG de 1943 a 2020.	32
Tabela 3 – Resultados dos testes (valor- p) de Ljung-Box e Mann- Kendall da série de precipitação máxima anual (mm), de 1943 a 1980, em Petrópolis-RJ, Teresópolis-RJ e Juiz de Fora-MG	32
Tabela 4 – Resultados dos critérios de convergência referente ao ajuste da distribuição GEV aos dados de precipitação máxima anual da cidade de Petrópolis-RJ (1943 a 1980), considerando as diferentes estruturas de priori.	33
Tabela 5 – Resultados do Critério de Desvio de Informação (DIC) , Erro médio de predição, Acurácia e Amplitude média.	34
Tabela 6 – Resultados (valor- p) dos testes de Ljung-Box e Mann- Kendall da série de precipitação máxima anual (mm), de 1943 a 2020, em Petrópolis-RJ, Teresópolis-RJ.	34
Tabela 7 – Resultados dos critérios de Convergência e do teste de aderência Aderson - Darling.	35
Tabela 8 – Estimativas (média a posteriori) e desvio padrão (DP) dos parâmetros da distribuição GEV ajustada aos dados de Petrópolis-RJ, com respectivos intervalos HPD com 95% e suas amplitudes, considerando a priori informativa baseada nas precipitações máximas de Teresópolis-RJ de 1943 a 2020.	35
Tabela 9 – Precipitações máximas preditas (mm) para diferentes tempos de retorno (em anos) em Petrópolis-RJ, com respectivos intervalos HPD de 95% de credibilidade e amplitude intervalar.	36
Tabela 10 – Estimativas dos parâmetros da distribuição GEV ajustada aos dados de Petrópolis-RJ via inferência bayesiana com respectivos intervalos HPD com 95% e amplitude intervalar sendo Juiz de Fora (MG) e Teresópolis-RJ as distribuições a priori, de 1943 a 1980.	42
Tabela 11 – Estimativas dos níveis de retorno da distribuição GEV ajustada aos dados de Petrópolis-RJ via inferência bayesiana com respectivos intervalos HPD com 95% , Erro médio de predição(EMP), Amplitude Intervalar média(Amp) e Acurácia para diferentes estruturas de distribuição a priori de 1943 a 1980.	43

SUMÁRIO

1	INTRODUÇÃO	12
2	REVISÃO DE LITERATURA	14
2.1	Análise da precipitação máxima	14
2.2	Distribuição generalizada de valores extremos (GEV)	17
2.3	Função de verossimilhança da distribuição GEV	18
2.4	Inferência Bayesiana	18
2.5	Distribuição a <i>Priori</i>	20
2.5.1	<i>Priori não Informativa</i>	20
2.5.2	<i>Priori Informativa</i>	20
2.5.3	<i>Distribuição a Posteriori</i>	21
2.6	Método de Monte Carlo via Cadeia de Markov	21
2.7	Teste de Anderson- Darling (AD)	22
2.8	Critério Desvio-Informação (DIC)	22
3	MATERIAL e MÉTODOS	24
3.1	Dados de Precipitação máxima de Petrópolis RJ	24
3.2	Análises Preliminares	24
3.2.1	<i>Análise Descritiva</i>	25
3.3	Função de Verossimilhança	25
3.4	Estimação via Inferência Bayesiana	25
3.5	Elicitação das Prioris	25
3.5.1	<i>Juiz de Fora-MG</i>	26
3.5.2	<i>Teresópolis-RJ</i>	27
3.6	Obtenção das distribuições a <i>posteriori</i>	28
3.7	Nível de Retorno	28
3.8	Avaliação das estruturas de priori	29
3.8.1	<i>Erro médio de predição</i>	29
3.8.2	<i>Acurácia</i>	30
3.8.3	<i>Amplitude média</i>	30
3.9	Recursos Computacionais	30
4	RESULTADOS E DISCUSSÃO	31
4.1	Análise descritiva e testes preliminares	31
5	CONCLUSÃO	37
	REFERÊNCIAS	38
A	TABELAS	42

B	PACOTES UTILIZADOS E ROTINA DOS CÓDIGOS:	43
----------	---	-----------

1 INTRODUÇÃO

As fortes chuvas ocorridas nos anos de 2011 e 2013 na região serrana do estado do Rio de Janeiro que provocaram um dos piores desastres do país (OTTERO; CHARGEL; HORA, 2018), foram a principal motivação para a escolha do tema desta pesquisa. Sendo assim, a análise estatística dos registros meteorológicos é fundamental para entender os padrões climáticos das cidades ou até mesmo regionais, que são base para o planejamento das atividades de agricultura, saúde, planejamento urbano, entre outras. A produção agropecuária, por exemplo, é influenciada pelo comportamento probabilístico das variáveis climáticas (precipitação, temperatura, vento, radiação solar), exigindo que decisões estratégicas de investimentos e decisões operacionais, como por exemplo a de irrigação, sejam tratadas em função dessas características climáticas.

Portanto, conhecer, analisar e prever um evento climático extremo, ao longo do tempo, é relevante para o planejamento de atividades vulneráveis a sua ocorrência. Neste estudo foi analisada a precipitação máxima anual (PMA) da cidade de Petrópolis-RJ, uma vez que, municípios erguidos em regiões de relevos e declividades proeminentes têm maior chance de sofrer os impactos de grandes volumes pluviométricos (TAVARES; FERREIRA, 2020). Dessa forma, para tal finalidade foi utilizada a distribuição Generalizada de Valores Extremos (GEV), apresentada por Jenkinson (1955), a qual expressa em uma única função os três tipos de distribuição assintótica de valores extremos (Gumbel, Fréchet e Weibull).

Existem alguns métodos utilizados para obter os estimadores dos parâmetros da distribuição GEV. Segundo Coles (2001), embora o método da máxima de verossimilhança seja um dos mais utilizados, a abordagem Bayesiana, por permitir que sejam incorporadas informações *a priori* para redução das incertezas acerca das estimativas, tem apresentado resultados mais acurado e precisos.

Apesar de diferentes trabalhos realizados analisarem a precipitação e suas consequências para diversas regiões utilizando métodos variados sobre as estimativas dos parâmetros da distribuição GEV, não foram encontrados estudos abrangendo o nível de retorno pluviométrico da cidade de Petrópolis-RJ, os quais aplicassem a inferência Bayesiana para a estimação dos parâmetros e para a predição da ocorrência de precipitações máximas.

Portanto, a realização deste estudo se mostra fundamental, uma vez que o excesso de água oriundo de eventos extremos e acentuado por condições deficientes de urbanização é importante para obtenção de informações sobre técnicas de previsão em determinadas regiões do Brasil tais qual a cidade objeto de estudo. Esse fato é de vital importância para se tentar diminuir e até mesmo evitar catástrofes no referido município, o qual é, também, um importante polo turístico, têxtil e de produção de

alimentos orgânicos do estado do Rio de Janeiro.

O objetivo principal deste trabalho é ajustar a distribuição GEV às séries históricas de precipitações máximas anuais da cidade de Petrópolis-RJ, para prever a ocorrência de precipitações máximas esperadas, através da abordagem Bayesiana, comparando diferentes estruturas de informações a *priori* (informativas e não informativas). Especificamente, pretende-se avaliar se as informações a respeito de precipitações máximas, de Juiz de Fora-MG e Teresópolis-RJ, que são cidades próximas a Petrópolis, favorecem o aumento da precisão nas previsões de precipitação máxima. A partir da escolha da melhor estrutura na especificação de *priori*, serão realizadas a previsão das precipitações máximas esperadas para os tempos de retorno de 2, 5, 10, 25, 50, 100, 115 e 135 anos.

2 REVISÃO DE LITERATURA

2.1 Análise da precipitação máxima

Prever a probabilidade de ocorrência de um evento hidrológico extremo, seja com relação às grandes precipitações que provocam enchentes, ou relacionados a grandes estiagens, é um fator fundamental para a avaliação e o planejamento de sistemas hídricos como reservatórios, canais, vertedores, bueiros, galerias de águas pluviais e outros (RIGHETTO, 1998).

Segundo Ayoade (2007) em meteorologia, o termo “precipitação” é usado para qualquer deposição em forma líquida ou sólida, derivada da atmosfera. Conseqüentemente, o termo refere-se às várias formas líquidas e congeladas de água, como a chuva (do latim, *pluvia*), neve, granizo, orvalho, geada e nevoeiro. Contudo, somente a chuva e a neve contribuem significativamente para os totais de precipitação, assim tem-se volumes diferentes de precipitação.

A precipitação máxima é definida como uma ocorrência extrema, com duração, distribuição temporal e espacial críticas para uma área (para uma bacia hidrográfica, estado ou região) com conseqüências imediatas sobre a agropecuária, a conservação do solo, de estradas, barragens e a drenagem, cujo dimensionamento adequado exige conhecimento de ocorrências extremas (ASSAD; MASUTOMO; ASSAD, 1992).

Precipitações pluviométricas máximas na região sudeste brasileiro ocorrem com maior probabilidade na primavera e no verão (CAVALCANTI *et al.*, 2009). Na primavera, há uma mudança no regime de chuvas e temperaturas na maior parte do Brasil.

Conforme Campos *et al.* (2009), o estudo das precipitações máximas é um dos caminhos para conhecer a vazão de projetos de estruturas hidráulicas que sempre são elaborados admitindo probabilidades de falha. Isto significa que podem ocorrer vazões maiores do que a vazão adotada no dimensionamento. Com isso, é importante se conhecer alguns tipos de Precipitação (QUADRO *et al.*, 1996).

Precipitações Convectivas são típicas das regiões tropicais, como o Brasil. O aquecimento desigual da superfície terrestre provoca o aparecimento de camadas de ar com densidades diferentes, o que gera uma estratificação térmica da atmosfera em equilíbrio instável. Se esse equilíbrio, por qualquer motivo (vento, superaquecimento), for quebrado, provocará uma ascensão brusca e violenta do ar menos denso, capaz de atingir grandes altitudes. As precipitações convectivas são de grande intensidade e curta duração, concentradas em pequenas áreas (chuvas de verão), além de serem importantes para projetos em pequenas bacias. Com o aumento da temperatura do ar sobre o continente, as chuvas provenientes desta faixa são acompanhadas por trovoadas e rajadas de vento. Particularmente as Regiões mais afetadas são: Sul, Sudeste e Centro-Oeste do País (ROSS, 1996).

Os maiores totais acumulados de chuva concentram-se principalmente nas Regiões Sudeste, Centro-Oeste e Amazônica. Na região Amazônica, segundo o estudo de Farias *et al.* (2017), o volume médio acumulado com valores superiores a 2307,4mm anual. Estas chuvas podem estar associadas à passagem de sistemas frontais e à formação do sistema meteorológico conhecido por Zona de Convergência do Atlântico Sul (ZCAS), cuja principal característica é a ocorrência de chuvas por vários dias, resultando em enchentes e deslizamentos de terra (CAVALCANTI *et al.*, 2009).

Segundo Conti, Furlan e Scarlato (1998), a intensidade (ou taxa) da precipitação é qualificada da seguinte forma:

- Chuva fraca: quando a intensidade é inferior a 2,5 milímetros por hora (mm/h);
- Chuva moderada: quando a intensidade é igual ou superior a 2,5 mm/h mas inferior a 10 mm/h;
- Chuva forte: quando a intensidade é igual ou superior a 10 mm/h mas inferior a 50 mm/h;
- Chuva violenta: quando a intensidade é superior a 50 mm/h (geralmente sob a forma de chuvaradas).

O conhecimento dessa caracterização sobre precipitação se faz necessário para que seja realizada a redução do grau de vulnerabilidade das comunidades frente aos desastres naturais. Segundo Dereczynski, Calado e Barros (2017), a Defesa Civil emite alertas à partir de 20 a 30 mm/h ou 100 mm acumulados em 24h, sendo que $1mm = 1l/m^2$. Assim, permite-se que a sociedade e os governos enfrentem os presumíveis efeitos desastrosos destes acontecimentos na população e no meio ambiente. Segundo MARENGO *et al.* (2009) na construção de sistemas de escoamento de águas pluviais, a probabilidade de falha admitida pode ser maior ou menor, dependendo do tipo de estrutura. Sendo assim, o conhecimento destas probabilidades de ocorrência de valores extremos é fundamental para a diminuição dos prejuízos sócio econômicos.

Pela necessidade de se determinar a probabilidade de ocorrência de valores para máximos ou mínimos surge a Teoria dos Valores Extremos (TVE), a qual é um ramo da estatística e da probabilidade que estuda o comportamento estocástico de eventos extremos associados a um conjunto de variáveis aleatórias (ou vetores aleatórios) com distribuição comum F . Dentro da denominação geral de extremos, incluímos o máximo e o mínimo, os quais empregam estatísticas de ordem de extremos ou excessos acima (ou abaixo) de limites altos (ou baixos). O importante é que as características e propriedades de distribuições desses extremos aleatórios são determinados pelas

caudas extremas (inferior e superior, no caso univariado) da distribuição subjacente F (MENDES, 2004).

Sendo assim, restringir a atenção às caudas de uma distribuição dará a vantagem de diversos modelos estatísticos adequados para as mesmas. Os modelos baseados em máximos coletados em blocos ou baseados em excessos de um limiar pré determinado, permitem implementar inferências mais precisas nas caudas e nos parâmetros desta distribuição F (MENDES, 2004).

Os princípios da *TVE*, inicialmente descritos por Fisher e Tippett (1928), determinam que a distribuição dos extremos convergem para o máximo de uma série de dados num determinado período. Essas distribuições tendem para uma das três distribuições, definidas à seguir (*I, II, III*), conhecidas, respectivamente, como Gumbel, Fréchet e Weibull (COLES, 2001).

- Distribuição Gumbel (Tipo I)

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

para $\mu \in \mathbb{R}$, $\sigma > 0$ e $-\infty < x < \infty$.

- Distribuição Fréchet (Tipo II)

$$G(x) = \exp \left\{ - \left(\frac{x - \mu}{\sigma} \right)^{-\xi} \right\}, \quad (2)$$

para $\mu \in \mathbb{R}$, ; $\sigma > 0$, $\xi > 0$ e $x > \mu$.

- Distribuição Weibull (Tipo III)

$$G(x) = \exp \left\{ - \left[- \left(\frac{(x - \mu)}{\sigma} \right)^{\xi} \right] \right\}, \quad (3)$$

para $\mu \in \mathbb{R}$, ; $\sigma > 0$, $\xi < 0$ e $x < \mu$.

Gnedenko (1943) apresentou as condições necessárias e suficientes para a convergência em distribuição do máximo para uma das três formas vistas nas equações 1, 2 e 3. (COLES, 2001; MENDES, 2004). Posteriormente, Jenkinson (1955) propôs a distribuição generalizada de valores extremos (GEV, do inglês, *Generalized Extreme Value*), que uniu essas três distribuições em uma única forma paramétrica.

2.2 Distribuição generalizada de valores extremos (GEV)

Os três tipos de distribuições, apresentadas na seção anterior, (I, II, III) podem ser vistos como membros de uma única família de distribuições: a distribuição acumulada GEV, a qual é a representação de três distribuições de valores extremos em uma única família de um único parâmetro ξ . A distribuição acumulada denotada por $F(x)$ é dada por:

$$F(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{1}{\xi}} \right\} \quad (4)$$

definida em, $-\infty < x < \mu - \frac{\sigma}{\xi}$, para $\xi < 0$, $\mu - \frac{\sigma}{\xi} < x < +\infty$ para $\xi > 0$, sendo μ , σ e ξ , respectivamente, os parâmetros de locação, escala e de forma com $\sigma > 0$.

As distribuições de valores extremos de Fréchet e de Weibull correspondem aos casos particulares em que (4) para $\xi > 0$ e $\xi < 0$, respectivamente. Quando o limite de $F(x)$ com $\xi \rightarrow 0$ obtém-se:

$$F(x) = \exp \left[- \exp \left(- \frac{x - \mu}{\sigma} \right) \right], \quad (5)$$

que é a função de distribuição acumulada Gumbel com parâmetros de escala e de locação μ e σ respectivamente, $\sigma > 0$. Da distribuição acumulada da GEV, obtém-se a f.d.p:

$$f(x) = \frac{1}{\sigma} \left\{ \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{1+\xi}{\xi}} \right\} \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}, \quad (6)$$

definida em, $-\infty < x < \mu - \frac{\sigma}{\xi}$ para $\xi < 0$, $\mu - \frac{\sigma}{\xi} < x < +\infty$ para $\xi > 0$. Como o limite de $f(x)$ com $\xi \rightarrow 0$, tem-se que:

$$f(x) = \frac{1}{\sigma} \exp \left[- \frac{x - \mu}{\sigma} - \exp \left(- \frac{x - \mu}{\sigma} \right) \right]. \quad (7)$$

Alguns métodos podem ser utilizados na estimação dos parâmetros de distribuições de valores extremos, entre eles destacam-se: métodos dos momentos (é método de estimação de cálculos mais simples, entretanto, os estimadores desse método são, em geral, menos eficientes do que os estimadores de máxima verossimilhança); o método da máxima verossimilhança (EMV), variações, tanto, dos métodos dos momentos quanto do método da máxima verossimilhança e a Inferência Bayesiana, que será apresentada.

2.3 Função de verossimilhança da distribuição GEV

Sejam X_1, X_2, \dots, X_n uma série de n variáveis aleatórias independentes e identicamente distribuídas, com distribuição GEV e $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ um conjunto de n observações amostrais. Assegurada a pressuposição de independência entre as n observações, tem-se que a função de verossimilhança da distribuição GEV é dada por:

$$L(\boldsymbol{\theta}|\mathbf{x}) = \frac{1}{\sigma^n} \prod_{i=1}^n \left\{ \left[1 + \xi \left(\frac{x_i - \mu}{\sigma} \right) \right]^{-\frac{(1+\xi)}{\xi}} \right\} \exp \left\{ \sum_{i=1}^n \left\{ - \left[1 + \xi \left(\frac{x_i - \mu}{\sigma} \right)^{-\frac{1}{\xi}} \right] \right\} \right\} \quad (8)$$

em que, $\boldsymbol{\theta} = (\mu, \sigma, \xi)$, são os parâmetros de localização, escala e forma respectivamente.

Na distribuição GEV os EMV de μ, σ, ξ , não possuem solução analítica e são obtidos numericamente pela maximização da função de verossimilhança. Embora esta dificuldade seja apresentada o grande uso dos EMV para a distribuição GEV se faz, devido ao fato de nos casos regulares os EMV serem consistentes e eficientes (sob o verdadeiro modelo), além de assintoticamente normais (MENDES, 2004).

Apesar destas boas propriedades e da predileção pelo método de máxima verossimilhança, por sua fácil compreensão metodológica, Smith (1985), ressalta que há problemas quanto a regularidade dos EMV em relação aos valores do parâmetro ξ :

- quando $\xi > -0,5$, os estimadores de máxima verossimilhança são completamente regulares;
- quando $-1 < \xi < -0,5$, os estimadores de máxima verossimilhança existem, mas são não regulares;
- quando $\xi < -1$, os estimadores de máxima verossimilhança não existem.

Uma alternativa para superar estas dificuldades, que tem apresentado bons resultados, é aplicação da inferência Bayesiana Paola *et al.* (2018).

2.4 Inferência Bayesiana

A teoria Bayesiana, como mencionado anteriormente surgiu em meados do século XVIII e foi desenvolvida por Thomas Bayes, que apresentou uma teoria subjetiva de probabilidade, fundamentada em um conhecimento *a priori* em associação a incertezas envolvidas no estudo. Na década de 1930, alguns estudos teóricos, como o desenvolvido por Jeffreys (1939), aplicados em problemas científicos, exigiam resoluções de integrais complexas, o que fez com que os métodos Bayesianos ficassem

mais limitados a problemas simples, sem ganhar espaço no meio científico. O problema dos cálculos das integrais foi resolvido na década de 1990 através do trabalho de Gelfand *et al.* (1990) que explorou um recurso de simulação dinâmica, o algoritmo Amostrador de Gibbs, implementado por Geman e Geman (1984), o qual faz uso da teoria das Cadeias de Markov. Os resultados alcançados nesse trabalho contribuíram com a ciência e, conseqüentemente, este algoritmo se estabeleceu como uma das ferramentas mais aplicadas na análise Bayesiana.

Segundo Gelman *et al.* (1995), a inferência Bayesiana é o processo que objetiva encontrar um modelo de probabilidade para um conjunto de dados e resumir o resultado por uma distribuição de probabilidade sobre os parâmetros do modelo e sobre quantidades não observadas, tais como a predição para novas observações. Aplicando-se a distribuição *a priori* $p(\theta)$ e a função de verossimilhança $L(\mathbf{x}|\theta)$, obtém-se a distribuição *a posteriori* de θ por meio do Teorema de Bayes, de modo que qualquer conclusão seja baseada nessa distribuição (BOX; TIAO, 2011). O teorema é denotado por:

$$p(\theta|\mathbf{x}) = \frac{L(\mathbf{x}|\theta)p(\theta)}{\int L(\mathbf{x}|\theta)p(\theta)d\theta}; \quad (9)$$

em que, $\mathbf{x} = (x_1, x_2, \dots, x_n)$ representa a amostra. E, além disso, como o denominador não depende de θ , tem-se que:

$$p(\theta|x) \propto L(\mathbf{x}|\theta)p(\theta) \quad (10)$$

Para extrair a informação de θ da distribuição *a posteriori* nas situações em que possui mais de um parâmetro é necessário, integrá-la. De acordo com Sorensen (1996) e Gamerman e Lopes (2006), a integração dessa distribuição *a posteriori* para obtenção de marginais, geralmente, não é analítica. Logo é necessário utilizar algoritmos iterativos para solucioná-la como o Metropolis Hastings e o Amostrador de Gibbs que são métodos fundamentados em Monte Carlo via Cadeias de Markov (MCMC). As técnicas MCMC possibilitam a geração de cadeias de Markov, que podem ser determinadas como processos descrevendo trajetórias, nas quais quantidades sucessivas são descritas probabilisticamente de acordo com o valor de seu predecessor imediato. Normalmente, estes processos tendem a um equilíbrio e as quantidades limites seguem uma distribuição probabilística que não varia mais, e é denominada, distribuição de equilíbrio (GAMERMAN; LOPES, 2006).

A análise Bayesiana permite que as incertezas sobre uma estimativa pontual sejam também descritas por um intervalo, o qual é chamado de intervalo de credibilidade, sendo construído com base na distribuição *a posteriori*.

De acordo com Paulino, Turkman e Murteira (2003), a região de credibilidade é definida numericamente e admite uma interpretação probabilística concreta e explícita,

sendo diferenciada da interpretação do intervalo de confiança da estatística clássica. Existe uma infinidade de regiões de credibilidade com o mesmo grau de credibilidade, e que, normalmente, se interessa em selecionar aquela que englobe todos os valores de θ com menor largura ou amplitude. Essa região é denominada como região de alta densidade *posteriori*, abreviadamente região *HPD* (do inglês, *Highest Probability Density*).

2.5 Distribuição a *Priori*

A distribuição a *priori* é uma (função de densidade(f.p)) ou função de densidade de probabilidade (f.d.p) atribuída para o parâmetro, antes de se observarem os valores amostrais ou experimentais. Suponha que se tenha um modelo estatístico com o parâmetro θ e na estatística Bayesiana o parâmetro é considerado aleatório e desta forma lhe é atribuída uma distribuição (DEGROOT; SCHERVISH, 2012).

A utilização de informação a *priori* em inferência Bayesiana requer a especificação de uma distribuição a *priori* para a quantidade de interesse θ . Esta distribuição deve representar (probabilisticamente) o conhecimento que se tem sobre θ antes da realização do experimento. Nas próximas seções, serão discutidas diferentes formas de especificação e tipos da distribuição a *priori* (PAULINO; TURKMAN; MURTEIRA, 2003).

2.5.1 *Priori não Informativa*

A *priori* não informativa informa os possíveis valores de θ como igualmente prováveis. Neste caso, fazendo $f(\theta) \propto k$ para θ variando em um subconjunto da reta significa que nenhum valor particular tem preferência (BERNARDO, 1994).

2.5.2 *Priori Informativa*

A distribuição a *priori* informativa assume algum grau de conhecimento acerca do parâmetro, ou seja, pondera quais valores de θ são mais prováveis segundo informações coletadas antes de se conhecer os dados (DEGROOT; SCHERVISH, 2012).

Neste trabalho, a distribuição a *priori* considerada foi a normal trivariada conforme proposta por Coles e Powell (1996),

$$f(\theta) \propto \exp \left\{ -\frac{1}{2}(\theta - \Phi_0)^t \Sigma_0^{-1}(\theta - \Phi_0) \right\} \quad (11)$$

em que, $\theta = (\mu, \log(\sigma), \xi)'$ é o vetor de parâmetros da distribuição GEV, $\Phi_0 = (\mu_0, \log(\sigma_0), \xi_0)$ o vetor de médias dos hiperparâmetros e Σ_0 a matriz de variâncias e covariâncias

(3 × 3) dada por:

$$\Sigma_0 = \begin{pmatrix} VAR(\mu_0) & cov(\log(\sigma_0), \mu_0) & cov(\xi_0, \mu_0) \\ cov(\mu_0, \log(\sigma_0)) & VAR(\log(\sigma_0)) & cov(\xi_0, \log(\sigma_0)) \\ cov(\mu_0, \xi_0) & cov(\log(\sigma_0), \xi_0) & VAR(\xi_0) \end{pmatrix}. \quad (12)$$

2.5.3 Distribuição a Posteriori

De acordo com DeGroot e Schervish (2012), a distribuição condicional do vetor paramétrico θ dado x_1, \dots, x_n é chamada de distribuição a *posteriori* de θ . Conforme visto na equação 10 e através do teorema de Bayes, a combinação de informações a *priori* sobre os parâmetros de forma independente e a função de verossimilhança, obtém-se a distribuição a *posteriori* do parâmetro θ , base da da estimação Bayesiana.

A distribuição conjunta a *posteriori* por não apresentar forma analítica, amostras da mesma devem ser obtidas numericamente, utilizando-se o método de Monte Carlo via Cadeia de Markov (Markov chain Monte Carlo - MCMC) o qual soluciona as dificuldades computacionais, como discutido a seguir.

2.6 Método de Monte Carlo via Cadeia de Markov

Os métodos de Monte Carlo via cadeias de Markov (MCMC) são uma alternativa aos métodos não iterativos em problemas complexos. Obtém-se uma amostra da distribuição a *posteriori*, calculando as estimativas amostrais desta distribuição. Para isso, utiliza-se técnicas de simulação iterativa, baseadas em cadeias de Markov, e assim os valores gerados não serão mais independentes. Os métodos de MCMC mais utilizados são, o amostrador de Gibbs e o algoritmo de Metropolis-Hastings. Simulando-se um passeio aleatório no espaço de θ que converge para uma distribuição estacionária. (GAMERMAN; LOPES, 2006; CASELLA; BERGER, 2011).

Uma cadeia de Markov é um processo estocástico $\{X_0, X_1, \dots\}$ tal que a distribuição de X_t dados todos os valores anteriores X_0, \dots, X_{t-1} depende apenas de X_{t-1} . Matematicamente, $F(X_t \in A | X_0, \dots, X_{t-1}) = F(X_t \in A | X_{t-1})$ para qualquer subconjunto A . Os métodos *MCMC* requerem ainda que a cadeia seja, homogênea, isto é, as probabilidades de transição de um estado para outro são invariantes; irreduzíveis, ou seja, cada estado pode ser atingido a partir de qualquer outro em um número finito de iterações; aperiódicas, caso não haja estados absorventes. e os algoritmos que serão vistos aqui satisfaçam a estas condições. A integral analisada é escrita, geralmente, como o valor esperado com respeito a alguma distribuição de probabilidade. Mais detalhes podem ser vistos em (DOBROW, 2016).

2.7 Teste de Anderson- Darling (AD)

Uma vez obtida estimativas dos parâmetros da distribuição GEV, faz-se necessário avaliar a aderência da distribuição à série de dados, sendo O teste de aderência de Anderson-Darling (AD), um dos mais recomendados (ABREU *et al.*, 2018). Este teste foi proposto pela primeira vez para reduzir erros no progresso estocástico por Darling (1957). De acordo com Naghettini e Pinto (2007), é um teste não-paramétrico que procura ponderar mais fortemente as caudas das distribuições, nas quais, as maiores (ou as menores) observações da amostra podem alterar sobremaneira a qualidade do ajuste.

O teste AD, baseia-se na diferença entre as funções de probabilidades acumuladas, empírica $F_N(x)$, e teórica, $F_X(x)$, de variáveis aleatórias contínuas. Entretanto, o teste AD dá mais peso às caudas, por meio da divisão das diferenças entre $F_N(x)$ e $F_X(x)$ por $\sqrt{F_x(x)[1 - F_x(x)]}$

Dada uma amostra aleatória X_1, X_2, \dots, X_n , associados a uma função de distribuição desconhecida $F_N(x)$. Admitindo-se uma específica função (teórica) de distribuição contínua $F_X(x)$, as hipóteses a serem testadas são:

$$\begin{cases} H_0 : F_N(x) = F_X(x) \\ H_1 : F_N(x) \neq F_X(x) \end{cases}$$

Como regra de decisão rejeita-se H_0 se o valor-p do teste for menor que o nível de significância (α) adotado.

2.8 Critério Desvio-Informação (DIC)

O critério *DIC* introduzido por Spiegelhalter *et al.* (2002) é, particularmente, usual nos problemas Bayesianos de seleção de modelos para os quais amostras da distribuição a *posteriori* dos parâmetros dos modelos foram obtidas por simulação de Monte Carlo em Cadeias de Markov (MCMC). Ele se trata de uma aproximação assintótica para amostras grandes, sendo válido quando a distribuição a *posteriori* é aproximadamente uma distribuição normal multivariada.

Define-se o desvio como:

$$D(\theta) = -2 \ln L(\theta) + C \quad (13)$$

em que θ é o vetor de parâmetros desconhecidos do modelo; $L(\theta)$ a função de verossimilhança e C é uma constante que se cancela e, portanto, não precisa ser conhecida na comparação de dois modelos. O critério *DIC* introduzido por (SPIEGELHALTER *et al.*, 2002) é dado por:

$$DIC = D(\hat{\theta}) + 2p_D \quad (14)$$

sendo $D(\hat{\theta})$ o desvio avaliado na média a posteriori e p_D é o número efetivo de parâmetros no modelo, que é dado por $p_D = D - D(\hat{\theta})$, em que $D(\hat{\theta}) = E[D(\hat{\theta})]$ é o desvio médio a posteriori que mede a qualidade do ajuste do modelo aos dados. Valores menores para DIC indicam melhores modelos, podendo estes serem negativos. Spiegelhalter *et al.* (2002) sugerem utilizar o seguinte critério para o módulo da diferença entre os valores de DIC de dois modelos, A e B, analisados. Esse critério é apresentado pela seguinte equação:

$$D = |DIC_A - DIC_B|. \quad (15)$$

Sua interpretação é:

- se $D < 5$, não se pode escolher o modelo somente por ter o menor DIC , logo, é necessário utilizar outros métodos;
- se $5 \leq D \leq 10$, considera-se que a diferença entre os modelos é substancial;
- se $D > 10$, definitivamente descarta-se o modelo com maior DIC .

3 MATERIAL e MÉTODOS

3.1 Dados de Precipitação máxima de Petrópolis RJ

Os dados utilizados foram obtidos junto a Agência Nacional de Águas (ANA), referente à cidade de Petrópolis na região serrana estado do Rio de Janeiro-RJ. Com um posto de coleta no Rio Cidade estação, nº 2243011, a qual possui 704 m de altitude. Desde janeiro de 1943 até dezembro de 2020 com 77 anos observados.

As observações de 1943 a 1980 foram usadas para estimar os parâmetros da distribuição GEV e calcular a precipitação máxima provável para os tempos de retorno de 10, 20, 30 e 40 anos. Das observações restantes, de 1981 a 2020, foram extraídas as precipitações máximas observadas em 10, 20, 30 e 40 anos, as quais foram usadas para verificar a acurácia e precisão das diferentes estruturas de *priori* na predição das precipitações máximas esperadas.

Tal região tem o clima tropical brando, com chuvas abundantes (2000 a 3000 mm anuais) especialmente de outubro a maio, sem estação seca, com temperatura média anual de 23° C. O município, também, é reconhecido por suas belezas naturais: um dos atrativos turísticos da região é a vegetação do tipo tropical leve - florestas tropicais (GUERRA; LOPES; FILHO, 2007).

Outra característica importante, é que a região, de acordo com Paulino (2005) tem grande relevância no estado do Rio de Janeiro na agricultura familiar e orgânica.

3.2 Análises Preliminares

Primeiramente, foi necessário verificar se as observações satisfazem à pressuposição de independência da função de verossimilhança. Com esse fim foi realizado o teste de Ljung e Box (1978) com um nível de 1% de significância e com as seguintes hipóteses:

$$\begin{cases} H_0 : & \text{A série de precipitação máxima anual é independente.} \\ H_1 : & \text{A série de precipitação máxima anual não é independente.} \end{cases}$$

Nas séries de precipitação máxima anual da cidade de Petrópolis foi realizada a análise de estacionariedade, ou seja, a verificação se esses dados possuem tendência ou não, para isso utilizou-se o teste de Mann-Kendall (MANN, 1945; KENDALL, 1975) ao nível de 5% de significância e com as seguintes hipóteses:

$$\begin{cases} H_0 : & \text{A série de precipitação máxima anual é estacionária.} \\ H_1 : & \text{A série de precipitação máxima anual não é estacionária.} \end{cases}$$

Para avaliar a aderência da distribuição GEV à série de dados, foi utilizado o teste de Anderson Darling, conforme (seção 2.8), uma vez que esse se mostra mais robusto

em eventos de natureza hidrológica (ABREU *et al.*, 2018)

3.2.1 Análise Descritiva

A análise descritiva dos dados de precipitação máxima anual, das localidades estudadas, foi baseada na estimação de medidas de posição (valores de mínimo e máximo, média e mediana) e dispersão (desvio padrão e coeficiente de variação).

3.3 Função de Verossimilhança

Para estimação dos parâmetros da distribuição GEV foram utilizados os métodos de máxima verossimilhança e inferência Bayesiana. As observações x_1, x_2, \dots, x_n de tamanho n em que assumisse independência entre elas, cuja função de densidade de probabilidade $f(\mathbf{x}|\boldsymbol{\theta})$, é dada por:

$$L(\boldsymbol{\theta}|\mathbf{x}) = \frac{1}{\sigma^n} \prod_{i=1}^n \left\{ \left[1 + \xi \left(\frac{x_i - \mu}{\sigma} \right)^{\frac{-(1+\xi)}{\xi}} \right] \right\} \exp \left\{ \sum_{i=1}^n - \left[1 + \xi \left(\frac{x_i - \mu}{\sigma} \right)^{-\frac{1}{\xi}} \right] \right\} \quad (16)$$

em que, $\boldsymbol{\theta} = (\mu, \sigma, \xi)$, são os parâmetros de localização, escala e forma respectivamente.

3.4 Estimação via Inferência Bayesiana

As informações para os hiperparâmetros foram extraídas das distribuições GEV, ajustadas via inferência Bayesiana para a precipitação máxima anual de Petrópolis-RJ, do período de 1943 a 2020, nos quais foram utilizados os pacotes EVD e EVDbayes.

Silva e Portela (2016) citam as vantagens da metodologia Bayesiana na Análise de Valores Extremos, entre as quais citam: combinar a informação contida na amostra, com outras informações, através da distribuição *a priori*; formalizar, quantificar e acompanhar a propagação de incertezas na estimação, auferir as estimativas preditivas da variável investigada e as incertezas provenientes do fenômeno da estimação dos parâmetros.

3.5 Elicitação das Prioris

Neste trabalho, a distribuição *a priori* considerada foi a normal trivariada conforme proposta por Coles e Powell (1996),

$$p(\boldsymbol{\theta}) \propto \exp \left\{ -\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\Phi}_0)^t \boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\theta} - \boldsymbol{\Phi}_0) \right\} \quad (17)$$

em que, $\boldsymbol{\theta} = (\mu, \sigma, \xi)$ é o vetor de parâmetros da distribuição GEV, $\boldsymbol{\Phi}_0 = (\mu_0, \log(\sigma_0), \xi_0)$ o vetor de médias dos hiperparâmetros e $\boldsymbol{\Sigma}_0$ a matriz de variâncias e covariâncias

(3 × 3) dada por:

$$\Sigma_0 = \begin{pmatrix} VAR(\mu_0) & cov(\log(\sigma_0), \mu_0) & cov(\xi_0, \mu_0) \\ cov(\mu_0, \log(\sigma_0)) & VAR(\log(\sigma_0)) & cov(\xi_0, \log(\sigma_0)) \\ cov(\mu_0, \xi_0) & cov(\log(\sigma_0), \xi_0) & VAR(\xi_0) \end{pmatrix}. \quad (18)$$

Fundamentalmente, uma *priori* deve apresentar as seguintes características: Respeitar o espaço paramétrico, refletir apropriadamente o conhecimento de especialistas. Conduzir à um processo de inferência com boas propriedades. É importante ressaltar que uma *priori* não, necessariamente, precisa ser uma função densidade de probabilidade (fdp) nem mesmo ser integrável.

Primeiramente foram consideradas as estimativas de verossimilhança para o chute inicial. Foi considerada uma *priori* não informativa, em que foram utilizados os hiperparâmetros:

$$\Phi_0 = (0; 0; 0) \text{ e } \Sigma_0 = \begin{bmatrix} 10^5 & 0 & 0 \\ 0 & 10^5 & 0 \\ 0 & 0 & 10^2 \end{bmatrix}.$$

Foram geradas cadeias com 320.000 iterações, sendo feito descarte (burn) 20.000 das primeiras iterações e realizados saltos (thin) a cada 60 primeiras iterações.

As informações para os hiperparâmetros foram obtidas ajustando-se a distribuição GEV aos dados de precipitações máximas de Juiz de Fora-MG e Teresópolis-RJ . A escolha dessas cidades se justifica pelo fato de possuírem uma série histórica de registros de mesmo tamanho, e por apresentarem características semelhantes tais como: clima, latitude e longitude.

3.5.1 Juiz de Fora-MG

Para elicitar a *priori* informativa foram utilizados dados de precipitação máxima da cidade de Juiz de Fora (MG), que está situada na região sudeste brasileira, próxima à cidade de Petrópolis-RJ como se observa na Figura 1. Neste trabalho analisou-se uma série histórica de máximos anuais de 77 anos da estação do INMET n° 83692 do período de janeiro de 1943 a dezembro de 2020, expressas em altura de lâmina d'água (mm).

De acordo com Toledo (2009) e o relatório do *Intergovernmental Panel on Climate Change* (IPCC) CHANGE *et al.* (2007) o município de Juiz de Fora está situado a 750 m acima do nível do mar. O clima nesta localidade é quente e temperado. O inverno é menos chuvoso do que o verão. Segundo a Köppen-Geiger a classificação do clima é CWA (Clima subtropical de inverno seco (com temperaturas inferiores a 18° C) e verão quente (com temperaturas superiores a 22° C), Ferreira, Batista e Vianna (2015). A pluviosidade média anual é de 1644 mm (MOTA *et al.*, 2008)

As observações de janeiro de 1943 a dezembro de 2020 da cidade de Juiz de Fora-MG, foram utilizadas para estimar os parâmetros da distribuição GEV.



Figura 1 – Mapa da Região Sudeste brasileiro.

Fonte: Elaborado a partir de base cartográfica do IBGE.

3.5.2 Teresópolis-RJ

A cidade de Teresópolis está localizada no topo da Serra dos Órgãos e está inserida no Domínio Tropical Atlântico, fazendo parte do sistema orográfico da Serra do Mar (AB'SABER; AB'SABER, 1977).

O clima de Teresópolis é definido como mesotérmico brando úmido do tipo Cwa (Köppen, 194; Nimer, 1979). A temperatura média anual foi de 19,80° C. Anualmente a precipitação média é de 1690,7mm, apresentando uma sazonalidade bastante definida, embora sem perda hídrica. A estação chuvosa estende-se de setembro até abril, quando as médias pluviométricas são, em sua maioria, superiores a 100mm. Nos demais meses do ano, a baixa pluviosidade caracteriza a estação seca da região. Nos meses de junho e agosto estes índices registraram valores inferiores a 60mm (MELLO *et al.*, 2001).

De acordo com Martins *et al.* (2018), para avaliar a flexibilização da informação a *priori*, foi avaliada uma estrutura de *priori* considerando: uma vez e quatro vezes a matriz de variâncias e covariâncias, Σ_0 .

As informações dos hiperparâmetros foram extraídas da distribuição GEV ajustada para a precipitação máxima das cidades de Teresópolis-RJ e Juiz de Fora-MG. O ajuste usando as informações como *priori* foram chamados de modelo PI-Tere-1var,

PI-Tere-4var, PI-JF-1var e PI-JF-4var e seus hiperparâmetros, variâncias e covariâncias são apresentados na Tabela 1, foram utilizados os pacotes EVD, EVDbayes e Coda que realiza as análises a posteriori.

Tabela 1 – Estimativas dos parâmetros da distribuição GEV via inferência Bayesiana e respectivas variâncias e covariâncias.

Priori	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$	$Var(\hat{\mu})$	$Var(\hat{\sigma})$	$Var(\hat{\xi})$	$Cov(\hat{\mu}, \hat{\sigma})$	$Cov(\hat{\mu}, \hat{\xi})$	$Cov(\hat{\sigma}, \hat{\xi})$
PNI	0	0	0	10000	10000	100	0	0	0
PI-tere-1var	64,1	16,58	-0,165	10,0	5,57	0,225	2,18	-0,240	-0,223
PI-tere-4var	64,1	16,58	-0,165	39,9	2,23	0,101	8,73	-0,959	-0,892
PI-JF-1var	64,6	20,5	-0,262	14,1	7,66	0,016	1,18	-0,200	-0,214
PI-JF-4var	64,6	20,5	-0,262	56,4	30,63	0,063	4,73	-0,800	-0,855

Fonte: Autor.

Depois foram ajustados os modelos nos quais foram utilizadas uma *priori* não informativa e quatro *prioris* informativas.

3.6 Obtenção das distribuições a *posteriori*

A partir das distribuições a *priori* propostas e da função de verossimilhança, foi obtida a distribuição a *posteriori* conjunta.

A distribuição conjunta a *posteriori* para a distribuição GEV é dada pelo produto de expressão (16) pela expressão (17) e por não apresentar uma forma analítica faz-se necessária a utilização do método MCMC por meio do algoritmo iterativo Metropolis-Hastings sendo geradas 120.000 amostras, realizando-se um descarte das 20.000 primeiras e um salto a cada 20 iterações.

A análise de convergência das cadeias foi realizada por meio de três critérios: Critério de Geweke (ZG), Critério de Raftery e Lewis e Critério de Heidelberger - Welch. No critério de Geweke (1992) se o valor da estatística for $|ZG| < 1,96$, indica que não há evidências para rejeitar que a média do início da cadeia é igual à média da parte final. Enquanto o critério de Raftery e Lewis (1992) foi verificado se o fator de Dependência é próximo de 1,0, mas de acordo com Gamerman e Lopes (2006) valores menores que 5 são aceitos. No critério de Heidelberger e Welch (1983), cuja hipótese nula avalia a estacionariedade da amostra gerada, esta hipótese não será rejeita o *valor-p* do teste for maior que o nível de significância adotado (5%).

3.7 Nível de Retorno

Coles (2001), o nível de retorno é uma estimativa probabilística da ocorrência de alguns destes eventos. É uma medida estatística tipicamente baseada em dados históricos que denotam o intervalo médio de recorrência durante um longo período de

tempo, e geralmente é usada para análise de risco (por exemplo, para decidir se um projeto deve ser autorizado a avançar em uma zona de certo risco, ou para projetar estruturas para suportar um evento com certo período de retorno). O termo é bastante utilizado na hidrologia e tem como definição o inverso da probabilidade de um evento ser igualado ou ultrapassado.

$$T = \frac{1}{p}, \text{ em que } T \text{ é dado em anos.}$$

Embora tenha sido amplamente adotado pela hidrologia, esse termo também pode ser utilizado por outras áreas do conhecimento, como a meteorologia e climatologia.

O nível de retorno (x_p), associado ao período de retorno T , é obtido a partir da solução da equação:

$$\int_{-\infty}^{x_p} f(x; \theta) dx = 1 - p, \quad (19)$$

em que: $p = \frac{1}{T}$, ou seja,

Ao aplicar a função inversa $p = \frac{1}{T}$, ou seja,

$$F(x_p) = (1 - p). \quad (20)$$

Ao aplicar a função inversa de F nos dois membros da equação (20), tem-se que:

$$\hat{x}_p = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left\{ 1 - [-\ln(1 - p)]^{-\hat{\xi}} \right\}, \quad (21)$$

para $\xi \neq 0$.

A estimativa \hat{x}_p do nível de retorno x_p para períodos de retorno $T = \frac{1}{p}$ é obtida substituindo as estimativas de μ , σ e ξ , na equação (20).

3.8 Avaliação das estruturas de priori

Para avaliar as estruturas de priori serão utilizados o Critério Desvio-Informação (DIC), a acurácia, o erro médio de predição e a amplitude média intervalar.

o Critério Desvio-Informação (DIC), introduzido por Spiegelhalter et al. (2002), foi utilizado de acordo com a seção (2.8).

3.8.1 Erro médio de predição

O erro médio de predição (EMP) é realizado dividindo-se a série observada em duas partes. A parte inicial dos valores observados é utilizado para o ajuste das distribuições GEV. Enquanto a segunda parte utiliza-se para a verificação do erro médio

de predição. Pode-se assim verificar o desempenho de cada inferência de previsões futuras. O erro médio de predição é dado por:

$$EMP = \frac{1}{k} \sum_{i=1}^k \left| \frac{x_i - \hat{x}_i}{x_i} \right| \times 100\%, \quad (22)$$

em que, x_i é o valor observado, \hat{x}_i é o valor extremo predito para o tempo de retorno i , usando-se a expressão (21) e k é o número de previsões.

3.8.2 Acurácia

Para avaliar a acurácia foi analisado se o valor observado da precipitação máxima para o i -ésimo tempo de retorno estava contido na estimativa intervalar de 95% de credibilidade.

3.8.3 Amplitude média

A amplitude de intervalo de confiança ou de credibilidade é dada pela diferença entre o limite superior (LS) e inferior (LI).

A amplitude intervalar média é dada pela média aritmética da amplitude intervalar:

$$\text{Amp média} = \frac{1}{k} \sum_{i=1}^k (LS - LI). \quad (23)$$

Para critério de decisão para seleção da estrutura de priori mais adequada para se realizar as predições de precipitação máxima, utilizou-se a seguinte ordem das métricas de avaliação: DIC, menor EMP, maior acurácia, e menor amplitude média intervalar.

3.9 Recursos Computacionais

Todas as análises estatísticas foram realizadas utilizando o software R (CORE, 2019) e seus respectivos pacotes: coda ((PLUMMER *et al.*, 2006), evd (STEPHENSON; RIBATET, 2008), evdbayes (STEPHENSON; RIBATET, 2014) e extRemes (GIL-LELAND, 2015).

Informações sobre os códigos R utilizados estão apresentadas no APÊNDICE B.

4 RESULTADOS E DISCUSSÃO

Serão analisados as análises descritivas das precipitações máximas anuais e os testes de independência e estacionariedade.

4.1 Análise descritiva e testes preliminares

Na Figura 2 foram exibidas as precipitações máximas anual de Petrópolis-RJ, Teresópolis-RJ e Juiz de Fora-MG, correspondente aos registros de 1943 à 2020. Pode-se observar que a cidade de Petrópolis apresenta uma maior ocorrência de precipitação máxima superior a 100mm.

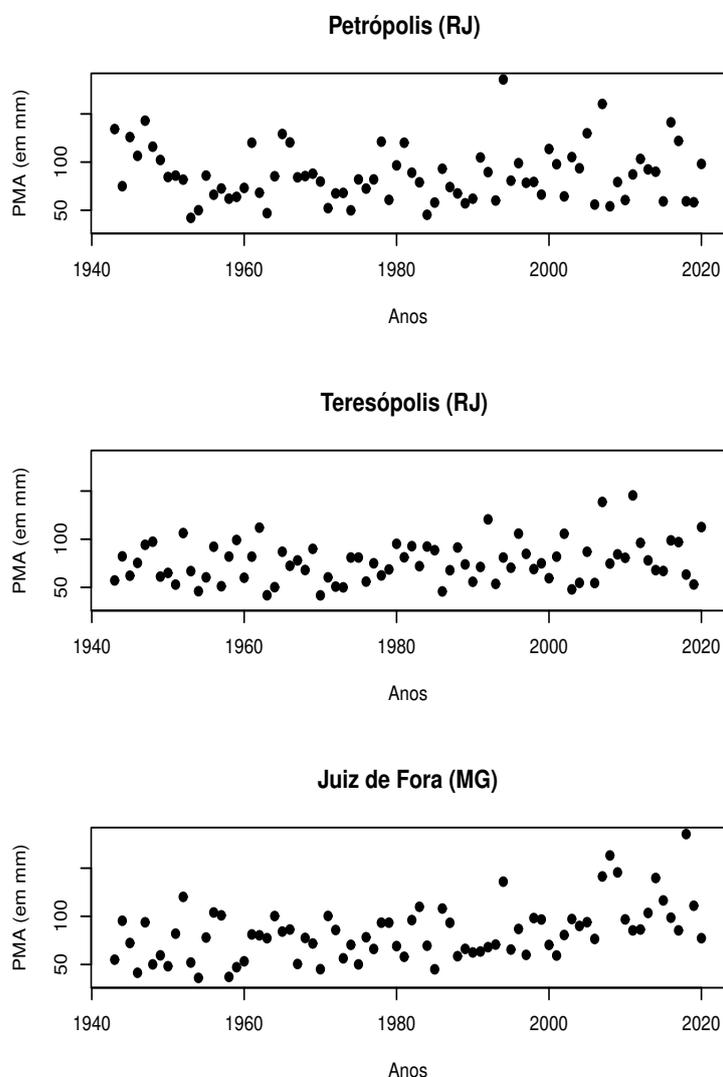


Figura 2 – Diagrama de dispersão da precipitação máxima anual (1943 à 2020) das cidades de Petrópolis-RJ, Teresópolis-RJ e Juiz de Fora-MG.

Fonte: Autor.

Na Tabela 2 foram apresentadas as estatísticas descritivas dos dados de precipitação máxima anual (1943 à 2020) das cidades de Petrópolis-RJ, Teresópolis-RJ e Juiz de Fora-MG. Observa-se que o terceiro quartil para os dados Petrópolis apresenta um valor superior a 100mm, indicando que das precipitações máxima anuais que ocorreram entre 1943 e 2020, 25% delas apresentaram valores maiores 101,4mm. Os valores máximos das precipitações máximas anual que ocorreram entre 1943 e 2020 nas cidades de Petrópolis-RJ, Teresópolis-RJ e Juiz de Fora-MG foram de 185,8, 145,5 e 185,4 mm, respectivamente.

Tabela 2 – Estatísticas descritivas da precipitação máxima anual (mm) de Petrópolis-RJ, Teresópolis-RJ e Juiz de Fora-MG de 1943 a 2020.

Localidade	Min	Q_1	Med	\bar{X}	Q_3	Máx	DP	CV (%)
Petrópolis	42,1	64,8	82,0	86,5	101,4	185,8	28,3	32,7
Teresópolis	41,7	60,4	74,9	76,4	89,7	145,5	21,3	27,9
Juiz de Fora	36,1	62,8	80,5	82,8	96,8	185,4	28,6	34,5

Fonte: Autor.

Observa-se também que dentre as três cidades, Juiz de Fora-MG é a que apresenta a menor precipitação máxima anual enquanto que as cidades de Petrópolis e Teresópolis, ambas do estado do Rio de Janeiro, apresentam valores próximos de menor precipitação máxima anual.

Verificou-se que as pressuposições de independência e estacionariedade das séries de observações com os testes de Ljung-Box e Mann-Kendall, respectivamente, sendo os resultados apresentados na Tabela 3.

Tabela 3 – Resultados dos testes (valor- p) de Ljung-Box e Mann- Kendall da série de precipitação máxima anual (mm), de 1943 a 1980, em Petrópolis-RJ, Teresópolis-RJ e Juiz de Fora-MG

Localidade	Ljung-Box	Mann-Kendall
Petrópolis	0,0403	0,1075
Teresópolis	0,4199	0,5974
Juiz de Fora	0,9995	0,6150

Fonte: Autor.

Pode-se concluir, ao nível de significância de 1%, que há independência e que não há presença de tendência, em todas as séries de dados (Tabela 3). Uma observação a ser feita, em relação à série de precipitação máxima anual de Petrópolis-RJ, é que esta apresentou um valor- p limiar ao nível de significância. Ao se observar o gráfico na Figura 2 poderia pensar em uma suspeita de presença de sazonalidade em parte da série de Petrópolis-RJ, porém, esta característica não será investigada no presente estudo. Desta forma, foi possível ajustar a distribuição GEV, via inferência considerando as diferentes estruturas de priori, aos dados de precipitação máxima anual da

cidade de Petrópolis, sendo os resultados dos critérios de convergência apresentados na Tabelas 4.

Tabela 4 – Resultados dos critérios de convergência referente ao ajuste da distribuição GEV aos dados de precipitação máxima anual da cidade de Petrópolis-RJ (1943 a 1980), considerando as diferentes estruturas de priori.

Prioris	Parâmetros	Raftery-Lewis	Geweke (ZG)	Heidelberger-Welch valor-p
PNI	μ	1,22	-0,24	0,588
	σ	1,02	-0,84	0,593
	ξ	1,07	1,19	0,654
PI-Tere-1var	μ	1,18	-0,99	0,184
	σ	1,06	0,23	0,925
	ξ	0,99	1,98	0,694
PI-Tere-4var	μ	1,52	0,72	0,846
	σ	1,09	0,75	0,390
	ξ	1,09	0,08	0,443
PI-JF-1var	μ	1,33	-1,77	0,632
	σ	1,11	1,67	0,416
	ξ	1,09	0,69	0,758
PI-JF-4var	μ	1,63	0,50	0,653
	σ	1,04	0,23	0,802
	ξ	1,11	-1,17	0,758

Fonte: Autor.

Observa-se, na Tabela 4, que para todos os parâmetros (μ, σ, ξ), o fator de dependência do critério de Raftery e Lewis estão próximos de 1, indicando independência entre as iterações. Pelo critério de Geweke, obteve-se que $|ZG| < 1,96$, designando que não há indícios de ausência de convergência. Pelo critério de Heidelberger e Welch, constatou-se que a série é estacionária ($p > 0,05$). Conclui-se então que não há evidências que indicam a não convergência das cadeias a posteriori para todos os parâmetros, considerando as diferentes estruturas de priori.

Uma vez em que não há evidências que indicam a não convergência das cadeias a posteriori, pôde-se avaliar quais distribuições a priori tem melhor desempenho na predição de precipitações máximas para Petrópolis. Para isso, foram calculadas as estimativas da precipitação máximas para os tempos de retorno 10, 20, 30 e 40 anos como respectivos intervalos HPD 95%, como pode ser observado nas Tabelas 10 e 11 no apêndice.

Posteriormente, analisou-se o DIC, o EMP, a acurácia e a amplitude média, conforme apresentado na Tabela 5. Pode-se observar que a diferença entre o maior valor de DIC (PI-Tere-1var) e o menor (PI-JF-4var), não foi maior que 5, sendo neste caso, de acordo com Spiegelhalter *et al.* (2002), não há uma diferença substancial entre os modelos, sendo então adotado como critério para a escolha da melhor estrutura de

priori, a que apresentou menor EMP. Deste modo, o modelo escolhido foi o que utiliza informações à *priori* fundamentado nos dados de precipitação máxima da cidade de Teresópolis-RJ, com os valores originais da matriz de variâncias e covariâncias (PI-Tere-1var), que, também, apresentou uma maior acurácia.

Tabela 5 – Resultados do Critério de Desvio de Informação (DIC) , Erro médio de predição, Acurácia e Amplitude média.

Prioris	DIC	EMP	Acurácia	Amp média
PNI	357,7	16,6	3 acertos	57,7
PI-Tere-1var	362,0	7,9	4 acertos	79,7
PI-Tere-4var	358,2	14,0	3 acertos	65,8
PI-JF-1var	358,4	12,1	3 acertos	56,7
PI-JF-4var	357,3	14,6	3 acertos	59,5

Fonte: Autor.

Uma observação a ser feita é que todas as estruturas de *priori* informativas apresentaram menor erro médio de predição de que a *priori* não informativa. Esses resultados corroboram com os obtidos por Coles e Powell (1996) e Martins *et al.* (2018), que obtiveram estimativas dos parâmetros da distribuição GEV e de precipitações máximas, para certos períodos de retorno, com maior acurácia e precisão, quando aplicaram a metodologia Bayesiana, destacando que ela oferece melhores resultados por permitir a utilização de informações à *priori*. Também, Lima, Kwon e Kim (2018) analisando os dados da Bacia Hidrográfica do Rio Han na Coreia do Sul, verificaram que a inferência Bayesiana apresentou bons resultados, com melhores estimativas da GEV e a redução das incertezas.

Para se realizar a predição das precipitações máximas esperadas para os tempos de retorno de 2, 5, 10, 25, 50, 100, 115 e 135 anos, em Petrópolis-RJ, considerando as informações advindas das precipitações máximas de Teresópolis-RJ, foi necessário verificar as pressuposições de independência e estacionariedade das séries de observações completas (1943 a 2020), de ambas as localidades. Os resultados dos testes de Ljung-Box e Mann-Kendall estão apresentados na Tabela 6.

Tabela 6 – Resultados (valor-*p*) dos testes de Ljung-Box e Mann- Kendall da série de precipitação máxima anual (mm), de 1943 a 2020, em Petrópolis-RJ, Teresópolis-RJ.

Localidade	Ljung-Box	Mann-Kendall
Petrópolis	0,8078	0,8799
Teresópolis	0,4815	0,1647

Fonte: Autor.

Observa-se, pelos resultados dos testes de Ljung-Box e Mann-Kendall ao nível de 1% de significância, que ambas as séries de dados são independentes e não apresentam tendência. Sendo assim, foram estimados os parâmetros da distribuição GEV, considerando a priori informativa baseada nas precipitações máximas de Teresópolis-RJ, foram avaliadas a convergência das cadeias a *posteriori* dos parâmetros, bem como a qualidade do ajuste da distribuição GEV à série de dados (Tabela 7).

Tabela 7 – Resultados dos critérios de Convergência e do teste de aderência Aderson - Darling.

Priori	Parâmetros	Raftery-Lewis	Geweke $ ZG $	Heidelberger-Welch	AD
PI_Tere_1var	μ	1,07	-1,44	0,067	0,3970
	σ	1,06	0,38	0,787	
	ξ	1,04	0,57	0,950	

Fonte: Autor.

Os resultados dos critérios de convergências podem ser observados na Tabela 7 em que os fatores de dependência de Raftery e Lewis estão próximos de 1, a estatística de Geweke $|ZG| < 1,96$, o valores- p do teste de Heidelberg-Welch foram não significativos ($p > 0,05$), indicando convergência das cadeias a *posteriori* dos parâmetros da distribuição GEV. Analisando-se o resultado do teste de Aderson - Darling, verificou-se que a distribuição GEV ajustou-se à série de dados ($p=0,3970$).

Na Tabela 8, são apresentadas as estimativas dos parâmetros com respectivos intervalos HPD de 95% de probabilidade e amplitude intervalar. Assim, posteriormente foram calculados os níveis de retorno para de 2, 5, 10, 25, 50, 100, 115 e 135 anos.

Tabela 8 – Estimativas (média a posteriori) e desvio padrão (DP) dos parâmetros da distribuição GEV ajustada aos dados de Petrópolis-RJ, com respectivos intervalos HPD com 95% e suas amplitudes, considerando a priori informativa baseada nas precipitações máximas de Teresópolis-RJ de 1943 a 2020.

Parâmetros	Estimativas	DP	Estimativas Intervalares		
			LI	LS	Amp
μ	65,1	1,7	61,9	68,3	6,4
σ	20,86	2,27	16,85	25,68	8,8
ξ	0,168	0,069	0,033	0,301	0,3

Fonte: Autor.

As estimativas obtidas no presente trabalho apresentaram valores próximos aos encontrados por Martins *et al.* (2018), embora analisando dados de outro estado, da cidade de Silvianópolis localizada no sul de Minas Gerais, de 1975 a 2016, obtiveram as seguintes estimativas dos parâmetros da GEV: $\hat{\mu} = 70,06$; $\hat{\sigma} = 25,4$ e $\hat{\xi} = 0,05$.

Na Tabela 9, são exibidos os níveis de retorno para 2, 5, 10, 25, 50, 100, 115 e 135 anos com a priori selecionada (PI_tere_1var) para o município de Petrópolis-RJ.

Tabela 9 – Precipitações máximas previstas (mm) para diferentes tempos de retorno (em anos) em Petrópolis-RJ, com respectivos intervalos HPD de 95% de credibilidade e amplitude intervalar.

Tempos de Retorno	Estimativas	Estimativas Intervalares		
		LI	LS	Amp
2	72,9	69,2	76,6	7,5
5	100,7	93,1	108,7	15,6
10	122,3	109,7	135,4	25,7
25	153,9	132,8	179,6	46,8
50	181,2	146,8	216,8	69,9
100	212,0	163,6	265,2	101,6
115	218,7	167,2	276,2	109,1
135	226,6	170,9	289,4	118,4

Fonte: Autor.

Uma interpretação prática para os níveis de retorno de precipitação máxima de Petrópolis-RJ, considerando o tempo de retorno de 50 anos (Tabela 9) é a seguinte: espera-se, em um tempo médio de 50 anos, que ocorra pelo menos um dia com precipitação máxima no ano, em Petrópolis-RJ, maior ou igual 181,2 mm, além disso, espera-se com 95%, de credibilidade que ocorra uma precipitação máxima maior ou igual a um valor que esteja entre 146,8mm e 216,8mm. Este resultado corrobora com os obtidos por Ottero, Chargel e Hora (2018), que utilizaram a distribuição Gumbel, a qual é um caso particular da GEV quando $\xi = 0$, para analisarem os dados de precipitação máxima das estações na região serrana do Rio de Janeiro para tempos de retorno de 5, 10, 20, 30, 50 e 100 anos. Os pesquisadores encontraram, para o tempo de retorno de 50 anos que, com 95% de confiança a precipitação máxima esperada estaria no intervalo de [99,4; 265,4] mm

Medeiros e Barros (2011) analisaram a mesma região e previram os níveis de retorno de máximas anuais para 10, 50 e 100 anos com o recurso da distribuição Gumbel e obtiveram [122; 157; 171] mm ratificando assim os resultados encontrados e apresentados na Tabela 9.

5 CONCLUSÃO

A distribuição GEV ajustou-se às séries de dados de precipitação máxima anual de Petrópolis-RJ e mostrou-se eficaz para realizar predições de precipitação máxima na localidade.

A série de precipitação máxima anual de Petrópolis-RJ não apresenta tendência.

A metodologia Bayesiana com a priori informativa baseada em informações da cidade de Teresópolis-RJ, considerando uma vez a variância, foi a mais adequada, permitindo prever a precipitação máxima para Petrópolis-RJ com maior precisão e acurácia, mostrando a eficiência da incorporação de conhecimentos a priori no estudo de precipitação máxima.

Na prática, espera-se que, em um tempo médio de 5 anos, ocorra pelo menos um dia com chuvas acima de 110,7 mm, nível acima do limiar considerado pela Defesa Civil para a emissão de alertas para a população.

Os valores de precipitação máxima preditos neste trabalho podem ser usados no planejamento de construções hidráulicas, para as atividades agrícolas e para as ações da Defesa Civil na uma vez que a cidade Petrópolis-RJ.

Referências

- ABREU, M. C. *et al.* Critérios para escolha de distribuições de probabilidades em estudos de eventos extremos de precipitação. *Revista Brasileira de Meteorologia*, SciELO Brasil, v. 33, p. 601–613, 2018.
- AB'SABER, A. N.; AB'SABER, A. Os domínios morfoclimáticos na América do Sul. primeira aproximação. 1977.
- ASSAD, E.; MASUTOMO, R.; ASSAD, M. Estimativa das precipitações máximas prováveis com duração de 24 horas e de 30 minutos. caso dos cerrados brasileiros. *Área de Informação da Sede-Artigo em periódico indexado (ALICE)*, Pesquisa Agropecuária Brasileira, Brasília, v. 27, n. 5, p. 677-686, maio. 1992, 1992.
- AYOADE, J. Introdução à climatologia para os trópicos. 11.ed. *Editora Bertrand Brasil*, Rio de Janeiro, 2007.
- BERNARDO, J. M. & *Adrian FM Smith: Bayesian Theory, fifth printing.* [S.l.]: Wiley, 1994.
- BOX, G. E.; TIAO, G. C. *Bayesian inference in statistical analysis.* [S.l.]: John Wiley & Sons, 2011. v. 40.
- CAMPOS, E. F. d. *et al.* Avaliação da distribuição da chuva nas vazões máximas urbanas usando dados de radar e de pluviógrafo. [sn], 2009.
- CASELLA, G.; BERGER, R. L. Inferência estatística. 2. ed. *São Paulo: Editora Cengage Learning*. 588p, 2011.
- CAVALCANTI, F. *et al.* *Tempo e Clima no Brasil*, São Paulo: *Oficina de Textos*, v. 1. [S.l.], 2009.
- CHANGE, O. C. *et al.* Intergovernmental panel on climate change. *World Meteorological Organization*, v. 52, 2007.
- COLES, S. *An introduction to statistical modeling of extreme values.* [S.l.]: Springer, 2001. v. 1.
- COLES, S. G.; POWELL, E. A. Bayesian methods in extreme value modelling: a review and new developments. *International Statistical Review/Revue Internationale de Statistique*, JSTOR, p. 119–136, 1996.
- CONTI, J. B.; FURLAN, S. A.; SCARLATO, F. *Clima eo meio ambiente.* [S.l.]: Atual, 1998.
- CORE, T. R. R: A language and environment for statistical computing. Vienna, Austria, 2019.
- DARLING, D. A. The kolmogorov-smirnov, cramer-von mises tests. *The Annals of Mathematical Statistics*, JSTOR, v. 28, n. 4, p. 823–838, 1957.

- DEGROOT, M. H.; SCHERVISH, M. J. *Probability and statistics, Fourth*. [S.l.]: Addison-Wesley, 2012.
- DERECZYNSKI, C. P.; CALADO, R. N.; BARROS, A. B. de. Chuvas extremas no município do rio de janeiro: Histórico a partir do século xix. *Anuário do Instituto de Geociências*, v. 40, n. 2, p. 17–30, 2017.
- DOBROW, R. P. *Introduction to stochastic processes with R*. [S.l.]: John Wiley & Sons, 2016.
- FARIAS, C. de S. *et al.* An analysis of the extreme rainfall case occurred in manaus during september 30th 2013. *Ciência e Natura*, Universidade Federal de Santa Maria-Centro de Ciências Naturais e Exatas, v. 39, n. 2, p. 436, 2017.
- FERREIRA, C. d. C. M.; BATISTA, G. M. F.; VIANNA, Y. C. G. Variações ou tendências climáticas: um estudo com as séries de temperatura do ar para a cidade de juiz de fora-mg. *Revista de Ciências Humanas*, n. 2, 2015.
- FISHER, R. A.; TIPPETT, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In: CAMBRIDGE UNIVERSITY PRESS. *Mathematical proceedings of the Cambridge philosophical society*. [S.l.], 1928. v. 24, n. 2, p. 180–190.
- GAMERMAN, D.; LOPES, H. F. *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. [S.l.]: CRC Press, 2006.
- GELFAND, A. E. *et al.* Illustration of bayesian inference in normal data models using gibbs sampling. *Journal of the American Statistical Association*, Taylor & Francis, v. 85, n. 412, p. 972–985, 1990.
- GELMAN, A. *et al.* *Bayesian data analysis*. [S.l.]: Chapman and Hall/CRC, 1995.
- GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, n. 6, p. 721–741, 1984.
- GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. *Bayesian statistics*, Clarendon Press, v. 4, p. 641–649, 1992.
- GILLELAND, E. *extremes: extreme value analysis*. *R Statistics*, 2015.
- GNEDENKO, B. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, JSTOR, p. 423–453, 1943.
- GUERRA, A. J. T.; LOPES, P. B. M.; FILHO, R. D. dos S. Características geográficas e geomorfológicas da apa, petrópolis, rj. *Revista Brasileira de Geomorfologia*, v. 8, n. 1, 2007.
- HEIDELBERGER, P.; WELCH, P. D. Simulation run length control in the presence of an initial transient. *Operations Research*, INFORMS, v. 31, n. 6, p. 1109–1144, 1983.
- JEFFREYS, H. *The theory of probability (1st)* Oxford. [S.l.]: England: Clarendon Press, 1939.

- JENKINSON, A. F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, Wiley Online Library, v. 81, n. 348, p. 158–171, 1955.
- KENDALL, K. Thin-film peeling-the elastic term. *Journal of Physics D: Applied Physics*, IOP Publishing, v. 8, n. 13, p. 1449, 1975.
- LIMA, C. H.; KWON, H.-H.; KIM, Y.-T. A local-regional scaling-invariant bayesian gev model for estimating rainfall idf curves in a future climate. *Journal of Hydrology*, Elsevier, v. 566, p. 73–88, 2018.
- LJUNG, G. M.; BOX, G. E. On a measure of lack of fit in time series models. *Biometrika*, Oxford University Press, v. 65, n. 2, p. 297–303, 1978.
- MANN, H. B. Nonparametric tests against trend. *Econometrica: Journal of the econometric society*, JSTOR, p. 245–259, 1945.
- MARENGO, J. A. *et al.* Mudanças climáticas e eventos extremos no brasil. *Rio de Janeiro: FBDS*, 2009.
- MARTINS, T. B. *et al.* Predição da precipitação máxima no município de silvianópolis-mg: Abordagens clássica e bayesiana. *IRRIGA*, v. 23, n. 3, p. 467–479, 2018.
- MEDEIROS, V. S.; BARROS, M. T. L. d. Análise de eventos críticos de precipitação ocorridos na região serrana do estado do rio de janeiro nos dias 11 e 12 de janeiro de 2011. *Anais do XIX Simpósio Brasileiro de Recursos Hídricos*, 2011.
- MELLO, C. d. *et al.* Análise de modelos matemáticos aplicados ao estudo de chuvas intensas. *Revista brasileira de ciência do solo*, SciELO Brasil, v. 25, p. 693–698, 2001.
- MENDES, B. V. M. *Introdução à análise de eventos extremos*. [S.l.]: Editora E-papers, 2004.
- MOTA, V. C. *et al.* Dependência temporal da precipitação pluvial e da temperatura do ar em juiz de fora, mg, por meio de análises geoestatísticas. *Revista Brasileira de Agrometeorologia*, v. 16, n. 3, p. 249–258, 2008.
- NAGHETTINI, M.; PINTO, E. Hidrologia estatística, 484 pp. *Belo Horizonte, CPRM-Serviço Geológico do Brasil, Brasil*, 2007.
- OTTERO, C. R.; CHARGEL, L. T.; HORA, M. d. A. G. M. d. Análise de frequência dos dados pluviométricos observados em 2011 e 2013 na região serrana, estado do rio de janeiro. *Revista Brasileira de Meteorologia*, SciELO Brasil, v. 33, p. 131–139, 2018.
- PAOLA, F. D. *et al.* Gev parameter estimation and stationary vs. non-stationary analysis of extreme rainfall in african test cities. *Hydrology*, Multidisciplinary Digital Publishing Institute, v. 5, n. 2, p. 28, 2018.
- PAULINO, C. D. M.; TURKMAN, M. A. A.; MURTEIRA, B. *Estatística bayesiana*. Lisboa: Fundação Calouste Gulbenkian, 2003.
- PAULINO, S. M. As cidades e as serras. espaço e identidades sociais na construção da ruralidade. *Identidades sociais: ruralidades no Brasil contemporâneo*, p. 255–276, 2005.

- PLUMMER, M. *et al.* Coda: convergence diagnosis and output analysis for mcmc. *R news*, v. 6, n. 1, p. 7–11, 2006.
- QUADRO, M. F. *et al.* Climatologia de precipitação e temperatura. *Climanálise (Boletim de Monitoramento e Análise Climática), Especial de*, v. 10, 1996.
- RAFTERY, A. E.; LEWIS, S. M. [practical markov chain monte carlo]: comment: one long run with diagnostics: implementation strategies for markov chain monte carlo. *Statistical science*, JSTOR, v. 7, n. 4, p. 493–497, 1992.
- RIGHETTO, A. M. *Hidrologia e recursos hídricos*. [S.l.]: EESC/USP, 1998.
- ROSS, J. L. S. *Geografia do brasil*. [S.l.]: Edusp, 1996. v. 3.
- SILVA, A. T.; PORTELA, M. M. Esquema bayesiano para estimar a distribuição de precipitações máximas anuais com duração subdiária em portugal continental. *Recursos Hídricos*, v. 37, n. 1, 2016.
- SMITH, R. L. Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, Oxford University Press, v. 72, n. 1, p. 67–90, 1985.
- SORENSEN, D. Gibbs sampling in quantitative genetics. *Intern Rapport. Statens Husdyrbrugsforsog (Denmark)*. no. 82., 1996.
- SPIEGELHALTER, D. J. *et al.* Bayesian measures of model complexity and fit. *Journal of the royal statistical society: Series b (statistical methodology)*, Wiley Online Library, v. 64, n. 4, p. 583–639, 2002.
- STEPHENSON, A.; RIBATET, M. evdbayes: Bayesian analysis in extreme value theory. *R package version*, p. 1–0, 2008.
- STEPHENSON, A.; RIBATET, M. evdbayes: Bayesian analysis in extreme value theory. *R package version*, v. 1, 2014.
- TAVARES, C. d. M. G.; FERREIRA, C. d. C. M. A relação entre a orografia e os eventos extremos de precipitação para o município de petropolis-rj. *Revista Brasileira de Climatologia*, v. 26, 2020.
- TOLEDO, C. C. Mudanças climáticas e a concentração de precipitação em juiz de fora–mg. *Jornal Eletrônico Faculdade Vianna Júnior*, v. 1, n. 2, p. 13–13, 2009.

A Tabelas

Tabela 10 – Estimativas dos parâmetros da distribuição GEV ajustada aos dados de Petrópolis-RJ via inferência bayesiana com respectivos intervalos HPD com 95% e amplitude intervalar sendo Juiz de Fora (MG) e Teresópolis-RJ as distribuições a priori, de 1943 a 1980.

Priori	Parâ- metros	Estima- tivas	DP	Estimativas Intervalares		
				LI	LS	Amp
PNI	μ	73,4	4,33	64,3	81,3	16,9
	σ	23,40	3,47	16,9	30,2	13,3
	ξ	-0,060	0,152	-0,371	0,220	0,6
PI_Tere_1var	μ	64,2	2,4	59,3	68,8	9,6
	σ	21,43	3,29	15,73	28,33	12,6
	ξ	0,201	0,110	-0,013	0,416	0,4
PI_Tere_4var	μ	68,9	3,4	62,2	75,4	13,2
	σ	22,1	3,2	16,32	28,52	12,2
	ξ	0,052	0,142	-0,220	0,325	0,5
PI_JF_1var	μ	67,3	2,7	62,1	72,8	10,7
	σ	21,70	3,03	16,32	27,95	11,6
	ξ	0,118	0,087	-0,054	0,284	0,3
PI_JF_4var	μ	70,8	3,7	63,793	77,859	14,1
	σ	22,36	3,16	16,561	28,750	12,2
	ξ	0,023	0,125	-0,228	0,265	0,5

Fonte: Autor.

Tabela 11 – Estimativas dos níveis de retorno da distribuição GEV ajustada aos dados de Petrópolis-RJ via inferência bayesiana com respectivos intervalos HPD com 95% , Erro médio de predição(EMP), Amplitude Intervalar média(Amp) e Acurácia para diferentes estruturas de distribuição a priori de 1943 a 1980.

Prioris	Tr	Estima- tivas	LI	LS	Amp	Valores obs
PNI		123,1	106,9	143,6	36,7	
PI_tere_1var		125,7	106,2	149,5	43,2	
PI_tere_4var	10	122,2	105,0	142,4	37,4	120,1
PI_JF_1var		123,5	107,9	142,5	34,7	
PI_JF_4var		122,9	106,3	143,1	36,8	
PNI		138,4	116,2	169,3	53,1	
PI_tere_1var		152,7	120,3	191,0	70,7	
PI_tere_4var	20	141,5	115,9	174,8	59,0	185,8
PI_JF_1var		145,2	121,5	173,4	51,9	
PI_JF_4var		140,8	116,6	170,9	54,3	
PNI		147,3	122,1	187,8	65,7	
PI_tere_1var		170,4	128,7	222,1	93,3	
PI_tere_4var	30	153,3	120,8	197,0	76,2	185,8
PI_JF_1var		158,7	129,0	194,0	64,9	
PI_JF_4var		151,5	122,3	190,2	67,9	
PNI		153,5	123,2	198,6	75,4	
PI_tere_1var		184,0	133,1	244,8	111,7	
PI_tere_4var	40	162,0	123,8	214,3	90,5	185,8
PI_JF_1var		168,7	135,4	211,4	76,0	
PI_JF_4var		159,2	125,3	204,4	79,2	

Fonte: Autor.

B Pacotes utilizados e Rotina dos Códigos:

Descrição de alguns pacotes utilizados nesta pesquisa da seção 3.11.

- Coda: Fornece funções para resumir e traçar o saída de simulações de Markov Chain Monte Carlo (MCMC), como bem como testes de diagnóstico de convergência para o equilíbrio distribuição da cadeia de Markov;
- EVD: Funções para distribuições de Valores Extremos, e estende funções de simulação, distribuição, quantil e densidade para distribuições de valores extremos paramétricos univariados e multivariados e fornece funções de ajuste que calculam estimativas de máxima verossimilhança para modelos de máximos univariados e bivariados e para modelos de limiar univariados e bivariados;

- EVDbayes: Calcula quantis da distribuição GEV a partir de amostras armazenadas dentro de uma cadeia de Markov, correspondendo a probabilidades na cauda superior;

Códigos R utilizados

Todos os códigos R utilizados para testes, ajustes da distribuição e estimação via inferência Bayesiana estão disponíveis online no endereço:

<https://github.com/sandravcds78/GEV-petropolis/tree/main>.