AVALIAÇÃO DO PARÂMETRO DE FORMA DA DISTRIBUIÇÃO GEV PARA SÉRIES DE PRECIPITAÇÕES ANUAIS MÁXIMAS EM MINAS GERAIS

Flávio Henrique Carvalho Cardoso [1] *& Eber José de Andrade Pinto* [2 3]

Palavras-Chave – Teoria de Valores Extremos, Análise de Frequência, Variabilidade Espacial.

INTRODUÇÃO

A compreensão dos eventos extremos de precipitação em Minas Gerais (MG), especialmente em escala diária, é fundamental devido à concentração de estruturas de elevado risco em cabeceiras de bacias hidrográficas, como barragens de mineração. Critérios inadequados na estimativa de quantis podem subestimar cheias de projeto e elevar os riscos de falhas e desastres ambientais.

A análise de eventos extremos em hidrologia se apoia em fundamentos probabilísticos, para os quais a teoria de valores extremos oferece arcabouço teórico amplamente aceito. Nesse contexto, a distribuição Generalizada de Valores Extremos (GEV), cuja função acumulada de probabilidade é dada pela equação (1), constitui alternativa versátil para representar máximos anuais de precipitação, em razão da flexibilidade decorrente de seus três casos particulares. Dentre seus parâmetros (σ - escala, μ – posição, e ξ - forma) o de forma exerce papel determinante, pois governa a cauda superior da distribuição, definindo se a variável segue uma família de cauda pesada (Fréchet), leve (Gumbel) ou limitada (Weibull). Pequenas variações nesse parâmetro produzem grandes diferenças nos quantis estimados para longos períodos de retorno. No entanto, a literatura tem demonstrado dificuldade em explicar sua variabilidade espacial (Tyralis *et al*., 2019).

$$F(x) = exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \text{ para } 1+\xi\left(\frac{x-\mu}{\sigma}\right) > 0, \sigma > 0, \mu \in \mathbb{R}, \xi \in \mathbb{R} \quad (1)$$

Papalexiou e Koutsoyiannis (2013), em avaliação global do ξ para precipitações diárias, constataram que séries mais longas produzem estimativas mais elevadas, em média, fato também observado por Serinaldi e Kilsby (2014). Esse resultado levou à proposição de uma equação para correção das estimativas. O estudo também identificou predominância de estimativas superiores a zero, indicativas da família Fréchet, mas com comportamento distinto da média global nas estações brasileiras. Contudo, a baixa densidade de estações pluviométricas em algumas regiões, como o Brasil, pode introduzir vieses quando resultados globais são interpretados em contextos locais.

No âmbito nacional, o tema vem sendo retomado por Duval *et al*. (2024), que iniciaram uma avaliação em larga escala do parâmetro de forma da distribuição GEV ajustada a séries de precipitação máxima diária anual, empregando milhares de estações. Os resultados preliminares indicam valores médios inferiores aos obtidos globalmente nos estudos supracitados e sugere, inclusive, redução dos valores médios com o aumento do tamanho das séries, levantando dúvidas sobre a adequabilidade da aplicação direta de formulações de correção de viés no contexto brasileiro.

Especificamente em MG, não foram identificadas investigações que analisem de forma sistemática o comportamento do parâmetro de forma aplicado às chuvas máximas diárias. Dessa forma, o presente

1) Departamento de Engenharia Hidráulica e Recursos Hídricos, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil. flaviohccardoso@gmail.com
2) Departamento de Engenharia Hidráulica e Recursos Hídricos, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil. eber@ehr.ufmg.br
3) Serviço Geológico do Brasil (SGB), Brasil. eber.andrade@sgp.gov.br

trabalho teve por objetivo investigar a variabilidade espacial do ξ no estado, visando analisar o tipo predominante de cauda. Além disso, verificou-se a influência do tamanho das séries sobre os valores estimados, bem como a aplicabilidade de fatores de correção, como a equação proposta por Papalexiou e Koutsoyiannis (2013), ao contexto mineiro.

METODOLOGIA

A metodologia adotada foi subdividida em duas etapas. A primeira consistiu no levantamento das estações pluviométricas em MG e na avaliação da qualidade dos dados. Em seguida, realizou-se a avaliação estatística do comportamento do parâmetro de forma da distribuição GEV, considerando as séries de precipitação diária máxima por ano hidrológico (PDMAH) extraídas de cada estação.

Foram levantadas séries de precipitação diária constantes nas bases de dados da Agência Nacional de Águas e Saneamento Básico (ANA) e do Instituto Nacional de Meteorologia (INMET). Após triagem inicial, considerando tamanho mínimo de 30 anos e percentual de falhas inferior a 30%, restaram 582 estações pluviométricas distribuídas em MG. Também foram excluídas 361 medições classificadas como inconsistentes após verificação individual das 652 medições diárias superiores a 150 mm.

O ano hidrológico (AH) foi definido com base nos acumulados mensais médios móveis de seis meses, iniciando-se no período chuvoso, entre outubro e março, e terminando ao final da estiagem, entre abril e setembro. As séries PDMAH foram extraídas concomitantemente à verificação de consistência de cada AH, por meio de adaptação dos critérios de Papalexiou e Koutsoyiannis (2013). Foram excluídos AH com mais de 1/3 de falhas, com pelo menos 30 dias de falhas no período chuvoso, e sem máximos dentre os 60% maiores valores da série. Também foram realizados testes de hipótese, que levaram à exclusão de AH identificados como *outliers* baixos e de séries cuja hipótese nula foi rejeitada, ao nível de significância regional de 5%, nas avaliações de aleatoriedade, homogeneidade e tendência. Ao final, restaram 496 estações válidas.

A distribuição GEV foi ajustada a cada série PDMAH pelo método dos momentos-L (MML), proposto originalmente por Hosking (1990). As estimativas de ξ foram avaliadas por estatísticas descritivas, gráficos de Curtose-L versus Assimetria-L e ajuste da distribuição normal pelo método dos momentos.

A relação entre tamanho das séries e estatísticas descritivas das estimativas de ξ foi avaliada conforme Papalexiou e Koutsoyiannis (2013), comparando estatísticas descritivas por faixas de 10 anos. A variabilidade espacial foi analisada por mapa de quadrantes de 1,5° em latitude e longitude, com pelo menos três estações válidas, calculando-se a média de ξ em cada quadrante.

RESULTADOS

As estatísticas descritivas do ξ obtido nos ajustes das séries PDMAH são apresentadas na Tabela 1. Observa-se que média e mediana foram nulas até, pelo menos, a segunda casa decimal. Esse comportamento, em conjunto com o histograma e o diagrama de Curtose-L versus Assimetria-L, cujo ponto médio empírico (quadrado vermelho) coincide com o ponto teórico da distribuição Gumbel (triângulo verde), sugere predominância de caudas superiores leves nas distribuições associadas às séries PDMAH em MG. Ressalta-se, contudo, o elevado desvio padrão, compatível com a variabilidade amostral esperada para o estimador em função do tamanho das séries. A distribuição normal ajustada às estimativas, apresentada na Figura 1, apresentou média igual a 0,00 e desvio padrão igual a 0,11. Na Figura 1 também foram plotadas as distribuições normais ajustadas por Papalexiou e Koutsoyiannis (2013), antes e após a correção de viés pelo tamanho amostral, evidenciando predominância de caudas mais leves em MG em comparação ao comportamento médio

global. Essa diferença é condizente com a análise espacial de Papalexiou e Koutsoyiannis (2013), com as observações de Marra *et al*. (2023) e com os resultados preliminares de Duval *et al*. (2024).

Tabela 1 – Estatísticas obtidas para os parâmetros de forma da distribuição GEV

n	média	mediana	min	máx	Q1	Q3	dp	g
496	0,00	0,00	-0,32	0,52	-0,07	0,07	0,11	0,20

Sendo: n – número de estações; Q1 – primeiro quartil; Q3 – terceiro quartil; dp – desvio padrão; g – coeficiente de assimetria.

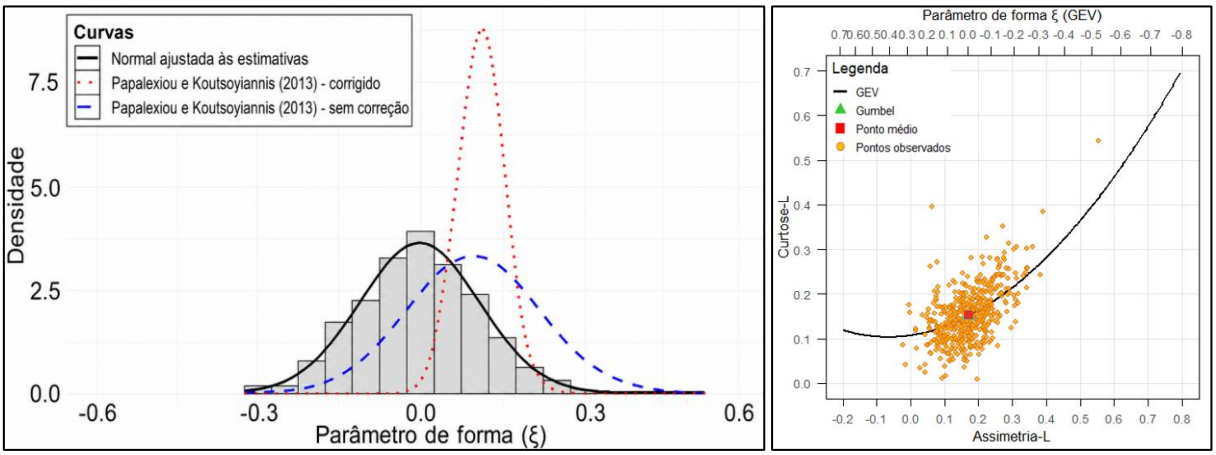


Figura 1. Histograma, ajuste da distribuição normal e gráfico Curtose-L x Assimetria-L. Fonte: autoria própria.

A Tabela 2 apresenta as estatísticas descritivas das estimativas do parâmetro de forma agrupadas por faixas de tamanho das séries. Os resultados não indicam influência sistemática do tamanho da série sobre os valores médios e medianos do parâmetro, que permaneceram próximos de zero, inferiores a 0,05, em todas as faixas analisadas. Conforme esperado, observou-se apenas redução gradual da dispersão, expressa pela diminuição do desvio padrão à medida que o tamanho das séries aumenta, comportamento compatível com a redução da variância amostral do estimador. Esses resultados diferem dos obtidos por Papalexiou e Koutsoyiannis (2013), sugerindo que, para MG, a aplicação da correção de viés proposta por esses autores poderia conduzir a resultados menos consistentes.

Tabela 2 – Estatísticas obtidas para as estimativas do parâmetro de forma para cada faixa de tamanho de série

Faixa	n	Média	dp	Mediana	Mínimo	Máximo	%>0 (Fréchet)	Q25	Q75	Q95	g
30-40	137	0,00	0,12	0,01	-0,31	0,28	56	-0,088	0,079	0,183	-0,25
41-50	132	0,00	0,12	-0,01	-0,32	0,52	45	-0,078	0,065	0,214	0,72
51-60	86	-0,02	0,10	-0,03	-0,26	0,20	41	-0,096	0,038	0,143	0,12
61-70	47	0,04	0,10	0,04	-0,21	0,31	62	-0,027	0,093	0,167	-0,07
71-80	78	0,00	0,08	0,00	-0,18	0,21	49	-0,058	0,053	0,114	0,16
81-90	16	0,00	0,09	-0,03	-0,15	0,15	44	-0,064	0,081	0,134	0,30

Sendo: n – número de estações; dp – desvio padrão; g – coeficiente de assimetria.

A Figura 2 apresenta a distribuição espacial das estações e a média do parâmetro de forma por quadrante em MG. Observou-se elevada dispersão dos valores estimados, com estações sugerindo as três famílias (Weibull, Gumbel e Fréchet) em praticamente todas as regiões do estado, sem padrões visuais de variabilidade espacial diretamente associáveis a características climatológicas ou geográficas regionais. Ainda assim, a avaliação das médias por quadrante indica predominância de valores próximos a zero, correspondentes a caudas leves.

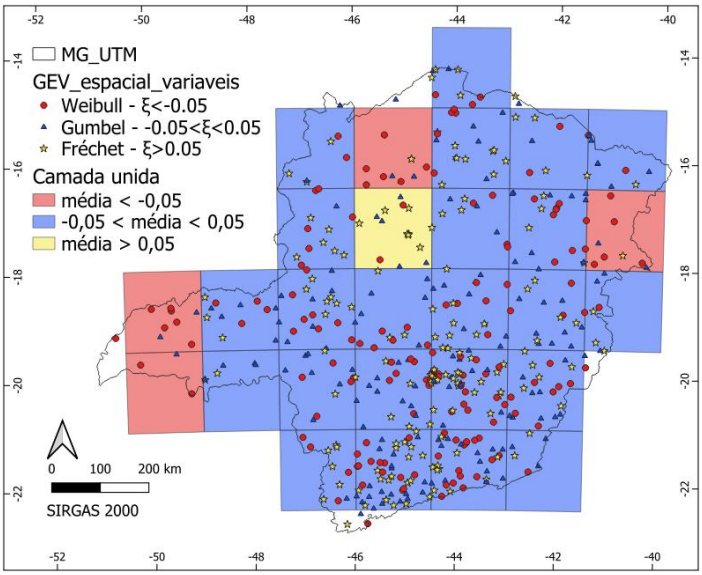


Figura 2. Variabilidade espacial do parâmetro de forma em MG. Fonte: autoria própria.

CONCLUSÕES

Esta pesquisa analisou o ξ da distribuição GEV aplicado às precipitações máximas anuais em MG. Os resultados indicaram predominância de estimativas próximas de zero com desvio padrão de 0,11, compatíveis com caudas superiores leves e distintas do comportamento médio global reportado na literatura. Não foi observada influência do tamanho das séries sobre os valores médios de ξ, sugerindo cautela na aplicação direta da correção de viés, como a proposta por Papalexiou e Koutsoyiannis (2013) em MG. Também não foram identificadas estruturas espaciais claras, sendo a dispersão observada compatível com a variabilidade amostral e com a incerteza da estimação do parâmetro.

REFERÊNCIAS

DUVAL, T. A. Q.; JUNIOR, D. S. R.; SOUZA, S. A. (2024). "*An assessment on the shape parameter of the Generalized Extreme Value distribution of annual daily maximum precipitation in Brazil.*" in Anais do 9th International Symposium on Integrated Water Resources Management (IWRM), 14th International Workshop on Statistical Hydrology (STAHY), I EBHE - Encontro Brasileiro de Hidrologia Estatística, Florianópolis, 2024. Resumo.

HOSKING, J. R. M. (1990). *"L-moments: Analysis and estimation of distributions using linear combinations of order statistics"*. Journal of the Royal Statistical Society. Series B (Methodological), v. 52, p. 105–124.

MARRA, F.; AMPONSAH, W.; PAPALEXIOU, S. M. (2023). *"Non-asymptotic Weibull tails explain the statistics of extreme daily precipitation".* Advances in Water Resources v. 173.

PAPALEXIOU, S. M.; KOUTSOYIANNIS, D. (2013). "*Battle of extreme value distributions: a global survey on extreme daily rainfall"*. Water Resources Research, v. 49, p. 187–201.

SERINALDI, F.; KILSBY, C. G. (2014). *"Rainfall extremes: Toward reconciliation after the battle of distributions".* Water Resources Research, v. 50, p. 336-352.

TYRALIS, H.; PAPACHARALAMPOUS, G.; TANTANEE, S. (2019). "*How to explain and predict the shape parameter of the generalized extreme value distribution of streamflow extremes using a big dataset*". Journal of Hydrology, v.574, p.628-645.