TÍTULO: COMPARAÇÃO ENTRE OS MÉTODOS DE MÁXIMA VEROSSIMILHANÇA E BAYESIANO PARA A ESTIMATIVA DE PRECIPITAÇÕES MÁXIMAS EM TEMPOS DE RETORNO LONGOS: ESTUDO DE CASO EM PASSO FUNDO, RS

Artur Vilaça Pereira[1]*; Bruna Rosário da Silva*[2]*; Veber Afonso Figueiredo Costa* [3]

Palavras-Chave – Valores extremos; Distribuição GEV; Inferência bayesiana; Precipitação máxima; Tempo de retorno.

INTRODUÇÃO

A estimativa de precipitações máximas associadas a tempos de retorno longos é insumo essencial ao dimensionamento de estruturas hidráulicas (vertedouros, galerias pluviais, bacias de detenção) e à gestão de riscos hidrológicos. A prática tradicional ajusta a distribuição Generalizada de Valores Extremos (GEV) por Máxima Verossimilhança (MLE), com intervalos de confiança obtidos pelo método delta. Essa abordagem, no entanto, depende de propriedades assintóticas que podem falhar em séries curtas, situação particularmente relevante diante das limitações de continuidade e consistência dos registros pluviométricos no Rio Grande do Sul (BRUBACHER; OLIVEIRA; GUASSELLI, 2020).

O parâmetro de forma da GEV, responsável por descrever o comportamento da cauda superior da distribuição (domínios de Gumbel, Fréchet ou Weibull), é historicamente o mais difícil de estimar com precisão em amostras pequenas, nas quais a curvatura da função de verossimilhança em torno do máximo é pouco pronunciada, o que pode gerar intervalos de confiança artificialmente estreitos, degenerados ou numericamente inválidos.

A inferência Bayesiana surge como alternativa a essa limitação, por não depender de aproximações assintóticas e por permitir incorporar informação prévia (priori) sobre os parâmetros da GEV, condensando a incerteza de estimação diretamente na distribuição a posteriori dos quantis de interesse, o que se torna uma vantagem relevante em contextos de dados escassos, como o observado em grande parte da rede brasileira.

Este trabalho tem como objetivo comparar as estimativas de precipitação máxima para tempos de retorno de 25 a 200 anos obtidas pelo método MLE tradicional e por uma abordagem Bayesiana com incorporação de informação regional. Ambos os métodos foram aplicados à distribuição GEV utilizando dados do posto pluviométrico de Passo Fundo, na região do Planalto Médio do Rio Grande do Sul. Para investigar o efeito do tamanho amostral sobre o desempenho dos dois métodos, o ajuste foi realizado para três subconjuntos amostrais (25, 50 e 71 anos), permitindo avaliar em que medida a incorporação de informação regional altera tanto as estimativas pontuais quanto a quantificação da incerteza associada aos eventos extremos.

METODOLOGIA

O estudo utilizou a série de máximas anuais de precipitação diária do posto pluviométrico de Passo Fundo, localizado na região do Planalto Médio do Rio Grande do Sul. Essa área possui clima subtropical úmido (Cfa, segundo a classificação de Köppen), com precipitação bem distribuída ao longo do ano e ocorrência frequente de eventos convectivos intensos associados a sistemas frontais. Foram utilizados três subconjuntos da série histórica, com 25, 50 e 71 observações anuais. O subconjunto de 10 anos foi testado em etapa preliminar e descartado por apresentar instabilidade

[1]) Programa de Pós-graduação em Saneamento, Meio Ambiente e Recursos Hídricos, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte/MG, artur.pereira@ufmg.br
[2]) Programa de Pós-graduação em Saneamento, Meio Ambiente e Recursos Hídricos, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte/MG, brs@ufmg.br
[3] Professor adjunto do Departamento de Engenharia Hidráulica e Recursos Hídricos da UFMG. E-mail: veber@ehr.ufmg.br

numérica no ajuste MLE, com intervalos de confiança degenerados pelo método delta, o que inviabilizaria uma comparação justa.

A distribuição GEV é caracterizada por três parâmetros: posição (μ), escala ($\sigma > 0$) e forma (ξ), sendo o quantil associado a um tempo de retorno T obtido pela inversão da função de distribuição acumulada. O ajuste por MLE foi realizado com a função fevd() do pacote extremes (GILLELAND; KATZ, 2016), em linguagem R, com intervalos de confiança de 95% construídos pelo método delta. O ajuste Bayesiano foi realizado com a mesma função fevd(), com 100.000 iterações de uma cadeia de Markov Monte Carlo (MCMC). Foram adotadas distribuições a priori normais e independentes para os parâmetros de posição, escala e forma da GEV, com hiperpârametros definidos a partir de informação regional. A convergência da cadeia foi avaliada por inspeção visual dos gráficos de trajetória (*trace plots*) dos parâmetros, verificando-se sua estabilidade ao longo das iterações.

A informação regional foi definida utilizando o estudo de Pickbrenner e Pinto (2025), sendo empregados os parâmetros de forma e escala do posto pluviométrico de Porto Alegre como referência para o ajuste Bayesiano. O posto de Porto Alegre foi adotado como fonte de informação regional devido à disponibilidade de parâmetros previamente estimados e documentados no Atlas Pluviométrico do Brasil (PICKBRENNER; PINTO, 2025), permitindo sua utilização como informação externa no ajuste Bayesiano. Os intervalos de credibilidade de 95% foram obtidos diretamente dos quantis 2,5% e 97,5% da distribuição a posteriori dos níveis de retorno. Esses intervalos incorporam a incerteza associada à estimação dos parâmetros da GEV, propagada para níveis de retorno por meio de suas distribuições a posteriori, condicionadas aos dados observados e às distribuições a priori adotadas.

Os dois métodos foram comparados quanto à magnitude das estimativas pontuais de nível de retorno para tempos de retorno de 25, 50, 100 e 200 anos. quanto à amplitude dos intervalos de incerteza e quanto à sensibilidade dessas propriedades ao tamanho amostral, avaliada por meio dos três cenários (25, 50 e 71 anos).

RESULTADOS

Em todos os cenários e tempos de retorno analisados, a abordagem Bayesiana com priori regional produziu estimativas pontuais sistematicamente superiores às do MLE tradicional (Tabela 1). Para T = 100 anos, por exemplo, a estimativa Bayesiana superou a do MLE em 24,4% no cenário de 25 anos (214,3 mm vs. 172,3 mm), 22,4% no cenário de 50 anos (189,3 mm vs. 154,7 mm) e 21,1% no cenário de 71 anos (181,1 mm vs. 149,5 mm). Esse padrão indica que a informação de forma da cauda extraída dos postos vizinhos é sistematicamente mais pesada do que aquela sugerida pela série local de Passo Fundo isoladamente.

Tabela 1 – Estimativas de precipitação máxima (mm) para T = 100 anos, por cenário e método.

Cenário	Método	Estimativa	Limite inf. 95%	Limite sup. 95%
25 anos	MLE	172,3	112,8	231,8
25 anos	Bayesiano	214,3	157,7	299,9
50 anos	MLE	154,7	127,9	181,5
50 anos	Bayesiano	189,3	152,3	242,0
71 anos	MLE	149,5	131,0	168,0
71 anos	Bayesiano	181,1	150,4	224,5

As estimativas dos parâmetros da GEV para cada método e tamanho amostral são apresentadas na Tabela 2, permitindo avaliar o efeito da incorporação da informação regional sobre os parâmetros da distribuição.

Tabela 2 – Estimativas dos parâmetros da GEV, por cenário e método.

Cenário	Método	μ (posição)	σ (escala)	ξ (forma)
25 anos	MLE	71.81485831	21.50408637	0.00670062
25 anos	Bayesiano	67.9312660	21.3646975	0.1359111
50 anos	MLE	76.09589126	19.31985920	-0.05445534
50 anos	Bayesiano	74.039831	20.0804392	0.1042576
71 anos	MLE	78.25358248	18.65548413	-0.08371892
71 anos	Bayesiano	76.31631698	19.37651630	0.07796071

Um segundo padrão relevante é que essa diferença relativa entre os métodos permanece praticamente estável à medida que o tamanho amostral aumenta de 25 para 71 anos (variando de 24,4% a 21,1% em T = 100), ao contrário do que ocorreria se a priori regional fosse apenas uma suposição fraca progressivamente "dominada" pelos dados locais. Essa persistência sugere que a informação regional carrega sinal genuíno sobre o comportamento da cauda e que a série local de Passo Fundo, mesmo com 71 anos, pode estar subestimando a real magnitude de eventos raros observados na região.

A Tabela 3 detalha essa diferença ao longo de todos os tempos de retorno, para o cenário mais robusto, de 71 anos. O afastamento entre os métodos cresce com o tempo de retorno, variando de 11,6% em T = 25 anos (145,8 mm vs. 130,6 mm) a 26,5% em T = 200 anos (199,9 mm vs. 158,1 mm). Além disso, os intervalos de credibilidade Bayesianos apresentaram maior amplitude que os intervalos de confiança do MLE. Ressalta-se, porém, que tais intervalos não são diretamente comparáveis, pois possuem construções e interpretações distintas: o intervalo de confiança está associado à variabilidade do procedimento de estimação, enquanto o intervalo de credibilidade expressa a incerteza a posteriori, condicionada aos dados e às prioris adotadas.

Tabela 3 – Estimativas de precipitação máxima (mm) por tempo de retorno, cenário de 71 anos.

T (anos)	MLE (mm)	IC 95% MLE	Bayesiano Reg. (mm)	ICred 95% Bayesiano Reg.
25	130,6	[118,9; 142,3]	145,8	[127,5; 169,9]
50	140,4	[125,5; 155,2]	163,1	[139,2; 195,8]
100	149,5	[131,0; 168,0]	181,1	[150,4; 224,5]
200	158,1	[135,5; 180,7]	199,9	[161,3; 256,3]

Nota-se ainda que, em ambos os métodos, a estimativa pontual para T = 100 anos diminui à medida que mais anos são incorporados à amostra. No MLE, os valores passam de 172,3 mm para 154,7 mm e depois para 149,5 mm, enquanto no método Bayesiano caem de 214,3 mm para 189,3 mm e chegam a 181,1 mm, considerando os cenários de 25, 50 e 71 anos, respectivamente. Esse efeito pode estar associado à regularização natural do parâmetro de forma quando eventos extremos isolados passam a representar uma fração menor da amostra total. Ainda assim, o método Bayesiano permanece consistentemente acima do MLE em todos os cenários, o que corrobora a hipótese de que a cauda regional da distribuição é de fato mais pesada do que a sugerida pela série local.

CONCLUSÕES

A incorporação de informação regional na abordagem Bayesiana resultou em estimativas de precipitação máxima sistematicamente superiores às obtidas pelo MLE tradicional que vai além da já conhecida robustez numérica em amostras pequenas: ao incorporar o parâmetro de forma estimado a partir de estação com parâmetro conhecido, o método Bayesiano passou a fornecer estimativas de

precipitação máxima sistematicamente mais altas e essa diferença não se dissipou com o aumento da série local, permanecendo estável mesmo no cenário de 71 anos. Isso indica que a cauda da distribuição de extremos na região do Planalto Médio gaúcho é mais pesada do que a própria série de Passo Fundo, isoladamente, seria capaz de revelar, sendo que o método Bayesiano regionalizado se mostra melhor porque compensa exatamente essa limitação estrutural de qualquer análise pontual baseada em um único posto, conseguindo reconhecer que o risco real de um local não se resume ao que já foi observado ali, mas ao que é fisicamente plausível na região como um todo.

Essa conclusão ganha relevância prática à luz dos eventos extremos que atingiram o Rio Grande do Sul em 2023 e, sobretudo, em maio de 2024, quando chuvas de magnitude sem precedentes no registro histórico recente causaram inundações catastróficas em dezenas de municípios, incluindo áreas do Planalto e do Vale do Taquari, próximas à região aqui estudada. Tais eventos ilustram, de forma concreta, o risco de subestimação inerente a abordagens baseadas exclusivamente no histórico local: séries de poucas décadas, por mais bem consistidas que sejam, podem simplesmente não conter, ainda, o evento mais extremo que a região é capaz de produzir. Uma abordagem que incorpora informação regional oferece uma margem de segurança mais compatível com o risco real observado, mesmo sem que se possa afirmar que o modelo aqui apresentado antecipasse especificamente aqueles eventos.

REFERÊNCIAS

GILLELAND, E.; KATZ, R. W. (2016). "extRemes 2.0: An Extreme Value Analysis Package in R". Journal of Statistical Software, v.72, n.8, p.1-39.

NAGHETTINI, M.; PINTO, É. J. A. (2007). Hidrologia Estatística. CPRM, Belo Horizonte/MG, 552 p.

PICKBRENNER, Karine; PINTO, Eber José de Andrade. Atlas pluviométrico do Brasil: equações intensidade-duração-frequência, estação pluviométrica: Porto Alegre - Jardim Botânico, códigos: 03051011 (ANA), 83967 e A801 (OMM/INMET), município: Porto Alegre, RS. Porto Alegre: SGB-Serviço Geológico do Brasil, 2025.

BRUBACHER, J. P.; OLIVEIRA, G. G.; GUASSELLI, L. A. (2020). “Preenchimento de Falhas em Séries Temporais de Precipitação Diária no Rio Grande do Sul”. Revista Brasileira de Meteorologia, v. 35, n. 2, p. 335-344. DOI: https://doi.org/10.1590/0102-7786352035