IMPROVING GEV SHAPE PARAMETER ESTIMATES FOR SUB-DAILY IDF CURVES UNDER SHORT RAINFALL RECORDS THROUGH HIERARCHICAL BAYESIAN INFERENCE

Larissa Zaira Rafael Rolim[1]; Tomás Antonio Quezado Duval[2]; Carlos Henrique Ribeiro Lima[3] & Dirceu Silveira Reis Junior[3]

Keywords – GEV distribution; precipitation extremes; Bayesian hierarchical models; short records.

1. INTRODUCTION

Intensity–duration–frequency (IDF) curves are fundamental tools for the design of urban drainage systems, hydraulic structures, and flood-control infrastructure, providing design rainfall estimates associated with different durations and return periods (Koutsoyiannis, Kozonis, Manetas, 1998). For sub-daily rainfall records, these estimates are particularly important because short and intense events are often associated with rapid hydrological responses, such as in urban flooding.

IDF curves are commonly derived by fitting frequency distributions to annual maximum rainfall series, with the Generalized Extreme Value (GEV) distribution widely used as a probabilistic basis for modeling block maxima (Naghettini, 2017). However, when sub-daily records are short, fragmented, or spatially limited, pure at-site estimation of the GEV parameters for each duration series can lead to large sampling uncertainty, especially under conventional estimation methods such as maximum likelihood and L-moments (Barbosa, Costa, 2023). This problem is particularly relevant for the GEV shape parameter, ξ, which controls the upper-tail behavior of the distribution and can strongly affect extrapolated return levels.

Bayesian hierarchical models provide a framework for sharing information across stations and durations, regularizing weakly informed parameters without completely suppressing local variability supported by the data (Lima, Kwon, Kim, 2018). In this context, the degree of pooling assigned to ξ may affect both tail flexibility and return-level uncertainty. This study evaluates different alternatives for modeling the GEV shape parameter under a Bayesian hierarchical approach for estimating sub-daily IDF relationships in the state Rio Grande do Sul, Brazil. The objective is to investigate how a regional Bayesian hierarchical model that assumes the shape parameter is independent of duration differs, in terms of return-level uncertainty in short sub-daily rainfall series, from one that allows the shape parameter to vary with duration. The Bayesian models are compared with conventional independent GEV fits based on maximum likelihood and L-moments.

2. METHODS

The study was conducted using sub-daily rainfall data from the State of Rio Grande do Sul, southern Brazil. Sub-daily annual maximum rainfall series were obtained from the AMP-RS dataset, compiled by Almeida *et al.* (2025) from records provided by different rainfall-monitoring operators in the state. Rainfall durations from 1 to 24 h were considered, and the dataset used in the study was constructed by applying temporal-consistency and data-availability filters. Annual maxima were extracted only from years in which missing or invalid rainfall aggregation windows did not exceed

1) Pesquisadora colaboradora, Universidade de Brasília. E-mail: larissazairarr@gmail.com
2) Mestrando, Universidade de Brasília. E-mail: tqduval@gmail.com
3) Professor Adjunto, Universidade de Brasília. E-mails: chrlima@unb.br; dirceureis@unb.br

20%, and only station–duration series with at least eight valid annual maxima were retained. The GEV distribution was used to represent the annual maximum rainfall intensity $X_{i,d,t}$, observed at station i, duration d, and year t, with location $\mu_{i,d}$, scale $\sigma_{i,d}$, and shape $\xi_{i,d}$ parameters. To represent the dependence of rainfall extremes on duration, a duration-dependent GEV formulation was adopted, in which the location and scale parameters vary systematically as functions of rainfall duration, following scaling structures used in Lima, Kwon, Kim (2018).

The focus of the analysis was to investigate the performance of different hierarchical structures for the GEV shape parameter ξ. The first Bayesian hierarchical model allows tail behavior to vary among stations, but assumes the shape parameter does not vary with duration, ξ_i, while the second model provides a greater flexibility allowing ξ to vary also with duration, $\xi_{i,d}$. In both cases, the hierarchical structure enabled information sharing across the estimated units, thereby reducing the instability associated with shape-parameter estimation in short station–duration series.

The Bayesian models were fitted using RStan with five chains of 10,000 iterations each, including 5,000 warm-up iterations. For the sake of comparison, independent at-site GEV models were fitted separately to each station–duration series using maximum likelihood (ML) and L-moments (L-MOM) estimators. In these cases, uncertainty was estimated using nonparametric bootstrap with 1,000 resamples per series.

Model evaluation considered predictive performance, parameter uncertainty, and return-level stability. The predictive performance of the Bayesian models was assessed using PSIS-LOO and WAIC (Vehtari, Gelman, Gabry, 2017). The GEV parameters μ, σ, and ξ were summarized using posterior credible intervals for the Bayesian models and bootstrap intervals for the conventional estimates. Return-level uncertainty was evaluated using the relative amplitude of the 95% interval for the 20-year return level, computed as the difference between the upper and lower limits of the interval divided by the median return level. For Bayesian models, these intervals correspond to posterior credible intervals, whereas for the conventional benchmarks they correspond to percentile bootstrap intervals.

3. RESULTS

The dataset for this study comprised 52 sub-daily rainfall stations with aggregated values for 24 durations (1-24 h), resulting in 1,248 station–duration series and 12,288 annual maximum observations. Valid record lengths ranged from 8 to 19 years, with most series concentrated between 8 and 10 years, characterizing a short-record setting in which independent at-site estimation of GEV parameters is expected to be highly uncertain, especially for the shape parameter. Among the Bayesian models, the pointwise predictive comparison favored the BHM station-specific ξ_i (independent of duration) structure according to both PSIS-LOO and WAIC, indicating that allowing the shape parameter to vary among stations improved predictive performance without requiring additional station-duration-specific flexibility.

The comparison of the relative amplitude of the 95% intervals showed a clear reduction in return-level uncertainty for the Bayesian hierarchical models. For the 20-year return level, the BHM station-specific ξ_i model resulted in a median 95% interval width of 16.4% of the median return level (Figure 1), while the BHM station-duration $\xi_{i,d}$ model resulted in a median relative width of 19.3%. These values were substantially lower than those obtained with the conventional independent GEV fits, which reached 35.7% for L-MOM and 266.3% for MLE. In addition to its much larger median uncertainty, the MLE benchmark showed an extremely wide interquartile range of relative widths, indicating severe instability in the estimation of extrapolated return levels under short-record conditions. The behavior of the GEV parameters was examined locally for station A801, which was selected because it had the longest valid record length in the dataset.

Figure 2 shows the estimates of the location, scale, and shape parameters as functions of rainfall duration. The location and scale parameters showed a similar decreasing pattern across methods, indicating consistent duration-dependent behavior for the central part of the distribution. Differences among methods were more evident for the shape parameter ξ, for which the L-MOM estimates displayed larger oscillations and wider bootstrap intervals, in some cases approaching or exceeding the $-0.5, +0.5$ range associated with the geophysical prior proposed by Martins, Stedinger (2000). In contrast, the Bayesian hierarchical models produced more regular estimates across durations, with stronger shrinkage toward a physically plausible range of tail behavior.

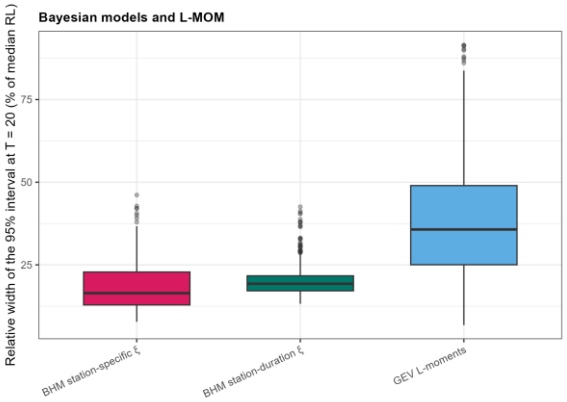


Figure 1. Relative width of the 95% uncertainty interval for 20-year return levels across station–duration series.

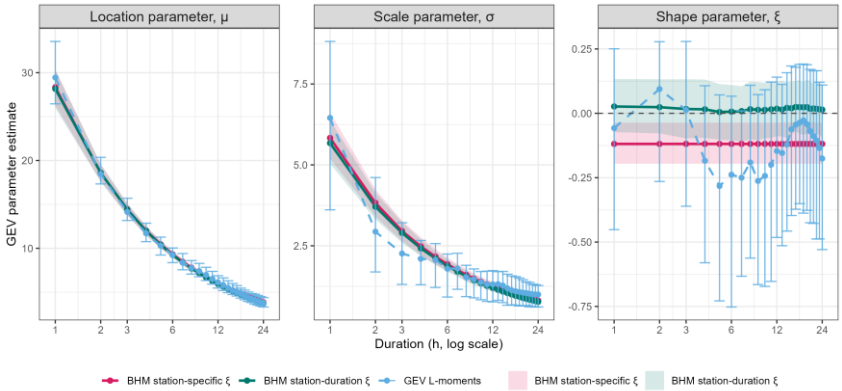


Figure 2. Duration-dependent estimates of GEV parameters at station A801.

The MLE estimates were not included in the parameter plot because their intervals were dominated by extreme values and, in some cases, by explosive parameter estimates. This behavior is consistent with Figure 1, in which the MLE model presented a median interval amplitude of 266.3%.

The local return-level comparison for station A801 is shown in Figure 3, using empirical plotting positions and fitted return-level curves for selected durations. The Bayesian hierarchical models reproduced the main trend of the observed annual maxima and produced relatively narrow uncertainty bands across durations. The L-MOM benchmark also followed the empirical pattern, but generally showed wider uncertainty bands, especially for shorter durations and larger return periods. The MLE benchmark, although retained in the return-level comparison for reference, showed broader and less stable intervals, reinforcing the sensitivity of independent maximum-likelihood fitting when applied to short station–duration series.

4. CONCLUSION

The findings indicate that reliable sub-daily IDF estimation from short records depends primarily on stabilizing the estimation of the GEV shape parameter. Hierarchical regularization reduces the influence of isolated extremes on tail estimates and provides more stable extrapolated

return levels than independent GEV fits. The results also suggest that greater local flexibility in ξ is not necessarily advantageous, since the station-specific hierarchical structure provided a better balance between predictive performance, uncertainty reduction, and hydrological plausibility. Therefore, Bayesian hierarchical modeling of the GEV shape parameter is a useful strategy for IDF estimation in regions with limited sub-daily rainfall records.

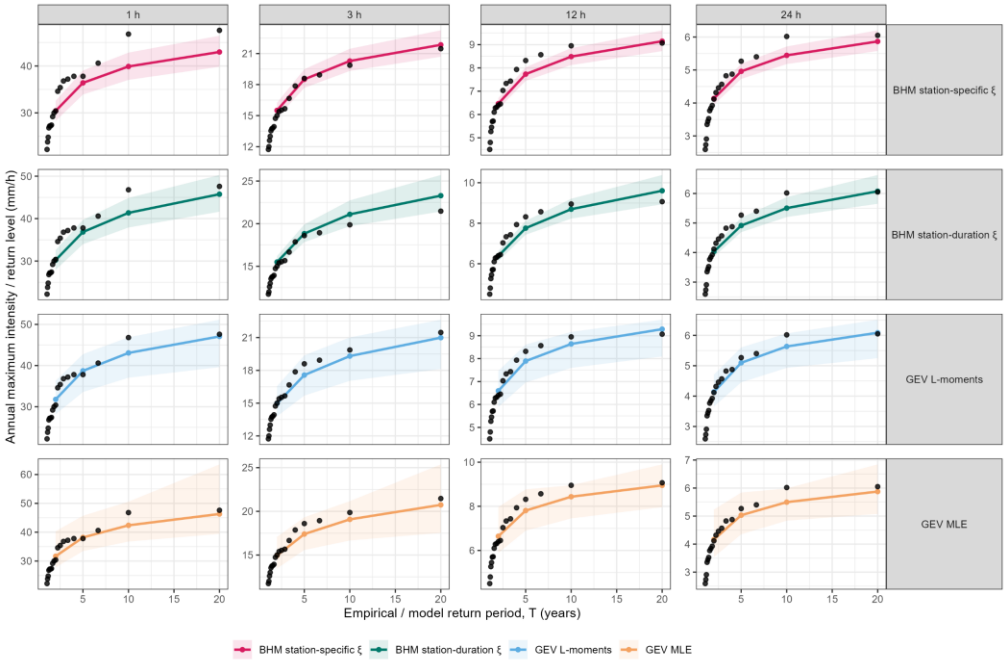


Figure 3. Empirical plotting positions and fitted return-level curves for station A801 at selected durations.

REFERENCES

ALMEIDA, C. D. N. et al. (2025). *"Avaliação de curvas Intensidade-Duração-Frequência (IDF) para o Rio Grande do Sul"*. Relatório técnico, João Pessoa/PB, 123 p. Available from: https://biblioteca.ana.gov.br/sophia_web/acervo/detalhe/111791.

BARBOSA, A. G.; COSTA, V. A. F. (2023). *"A review on the scaling properties in maximum rainfall marginal distributions: theoretical background, probabilistic modeling, and recent developments"*. Stochastic Environmental Research and Risk Assessment, v.37, n.12, p.4541–4553.

KOUTSOYIANNIS, D.; KOZONIS, D.; MANETAS, A. (1998). *"A mathematical framework for studying rainfall intensity-duration-frequency relationships"*. Journal of Hydrology, v.206, n.1, p.118–135.

LIMA, C. H. R.; KWON, H.-H.; KIM, Y.-T. (2018). *"A local-regional scaling-invariant Bayesian GEV model for estimating rainfall IDF curves in a future climate"*. Journal of Hydrology, v.566, p.73–88.

MARTINS, E. S.; STEDINGER, J. R. (2000). *"Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data"*. Water Resources Research, v. 36, n. 3, p. 737–744. DOI: 10.1029/1999WR900330.

NAGHETTINI, M. (2017). *Fundamentals of statistical hydrology*. Springer International Publishing, Cham, 660 p. DOI: https://doi.org/10.1007/978-3-319-43561-9.

VEHTARI, A.; GELMAN, A.; GABRY, J. (2017). *"Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC"*. Stat Comput, v. 27, p. 1413–1432. Available from: https://doi.org/10.1007/s11222-016-9696-4.

ACKNOWLEDGEMENTS

The authors acknowledge the National Water and Basic Sanitation Agency (ANA), under SEI Process No. 23106.017474/2024-93, and the Foundation for Scientific and Technological Enterprises (FINATEC) for the institutional, technical, and financial support that made the development of this work possible.