



Supplement of

Residual error modelling for process-based flood frequency analysis

Jonathan Romero-Cuellar et al.

Correspondence to: Jonathan Romero-Cuellar (jromeroc@uwaterloo.ca)

The copyright of individual parts of the supplement might differ from the article licence.

Contents of this file

Section S1. Introduction

Section S2. Modifications to CoSMoS for Flood Frequency Analysis

Figures S1

Section S3. Quantitative comparison of flood-quantile estimates

Table S1

Section S4. Diagnostic assessment of the GEV fits

Figure S2 to S7

S1. Introduction

This supporting information provides technical details on key modifications made to the CoSMoS framework to enhance its suitability for flood frequency analysis. Two primary adjustments were implemented to improve the realism and reliability of synthetic hydroclimatic time series used in uncertainty quantification:

1. Smoothing Monthly Transitions in Temperature Time Series

A bootstrap-based resampling method was introduced to mitigate discontinuities at month boundaries in temperature simulations. This approach preserves monthly statistical properties while ensuring seasonal continuity, which is critical for realistic hydrologic modeling.

2. Physical Constraints on Extreme Precipitation

To address biases in the representation of extreme precipitation events, quantile mapping was applied to align synthetic precipitation distributions with observed data. This correction improves the fidelity of upper-tail behavior, which is essential for accurate flood risk estimation.

Together, these enhancements support the generation of more realistic climate inputs for hydrologic models, thereby strengthening the reliability of flood quantile estimates produced by the UQ-flood framework.

S2. Modifications to CoSMoS for Flood Frequency Analysis

S2.1. Smoothing Monthly Transitions in Temperature Time Series

The CoSMoS framework is designed to simulate hydroclimatic time series by fitting monthly marginal distributions and autocorrelation structures independently. While this approach effectively captures intra-month statistical properties, it introduces abrupt discontinuities at month boundaries. These discontinuities are particularly evident in temperature time series (Tmax and Tmin), where the transition from one month to the next can result in unrealistic jumps in statistical behavior (see Figure S1, top panel).

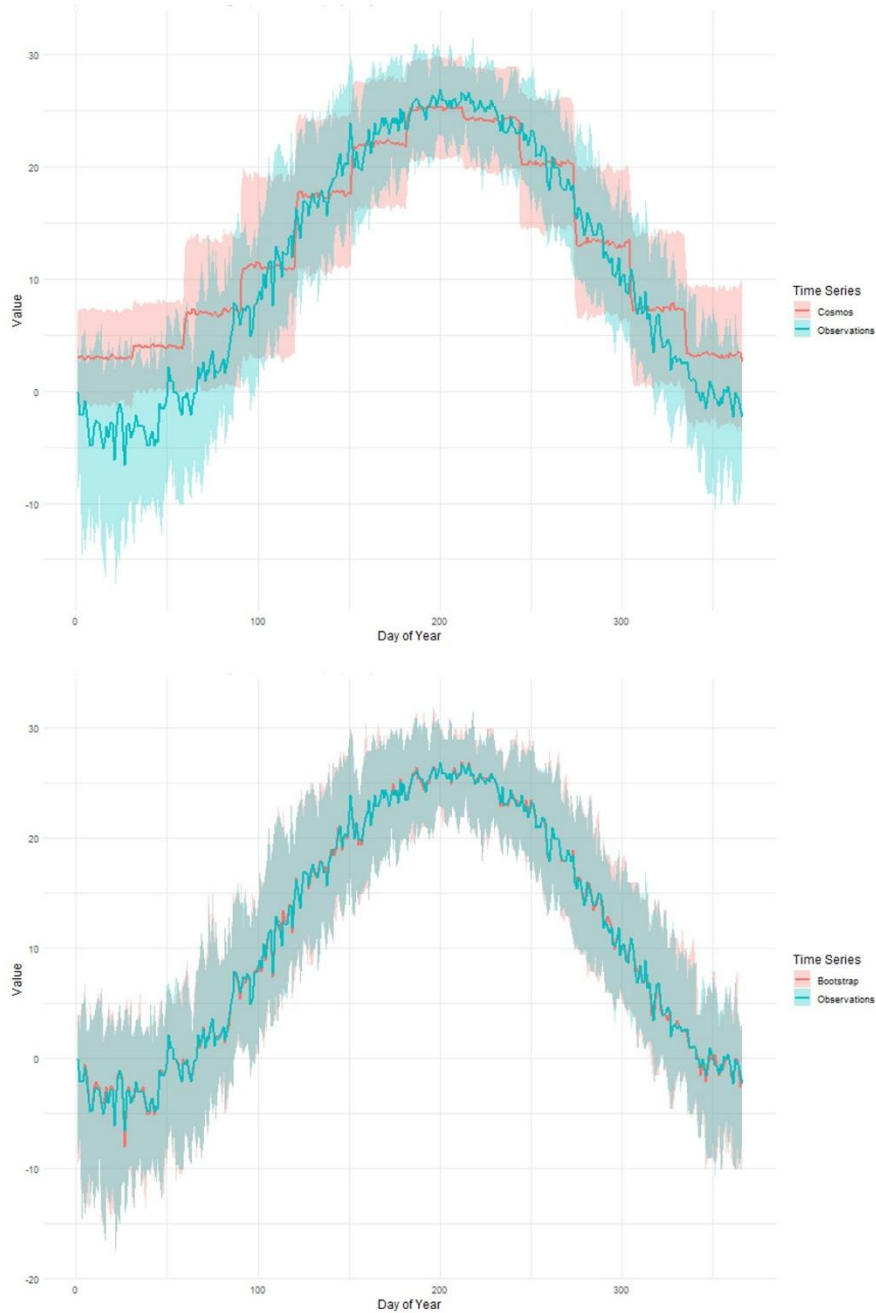


Figure S1. Comparison of daily maximum temperature time series across the water year: Observed data vs. CoSMoS (top) and Bootstrap-based synthetic data (bottom). Shaded areas represent interquartile ranges, highlighting the improved continuity and seasonal realism achieved through bootstrap resampling.

To mitigate this issue, we implemented a bootstrap-based smoothing strategy for temperature variables. The goal was to preserve the monthly statistical characteristics while ensuring continuity across month boundaries. To generate a 1000-year synthetic daily maximum temperature time series, we applied a non-parametric bootstrap approach based on resampling historical water years. This method preserves the seasonal structure, inter-month continuity, and temporal dependencies inherent in the observed data (Figure S2, bottom panel).

This modification enhances the realism of the synthetic temperature series, which is crucial for accurate simulation of hydrologic responses and flood frequency distributions.

S2.2. Physical Constraints on Extreme Precipitation

Although CoSMoS effectively simulates precipitation time series by fitting marginal distributions and autocorrelation structures, it tends to overfit the central portion of the distribution, particularly around the median (50th percentile), while underrepresenting the upper tail (e.g., 99th percentile). This imbalance can lead to synthetic precipitation values that are either unrealistically extreme or insufficiently representative of observed extremes, ultimately distorting flood frequency estimates when used as input to hydrologic models.

This limitation stems from CoSMoS's objective to match the overall distribution, which inherently places greater emphasis on the bulk of the data. As a result, the tail behavior—critical for flood risk analysis—is often misrepresented, producing synthetic events that exceed physically plausible limits or fail to capture observed extremes.

To address this issue, we applied quantile mapping (Cannon et al., 2015), a widely used bias adjustment technique in hydroclimatic modeling. This method aligns the distribution of synthetic precipitation with the observed distribution, with improvement in the representation of extremes. We constructed the empirical cumulative distribution functions (CDFs) of both the observed and CoSMoS-generated precipitation over the full historical record. Each synthetic value was then mapped to its corresponding quantile in the observed distribution, ensuring that the corrected series preserved the frequency and magnitude of observed extremes. This correction was applied globally across the time series to improve the realism of the upper tail, which is essential for reliable flood frequency analysis.

S3. Quantitative comparison of flood-quantile estimates

To complement the flood-frequency curves presented in Figs. 3 and 4 of the main manuscript, Table S1 provides a quantitative comparison of the central flood-quantile estimates for the Black River, LaHave River, and Moyie River watersheds. Estimates are reported for return periods of 2, 5, 10, 20, 50, and 100 years.

The comparison includes the empirical annual-maximum-flow estimates derived from the available long observational record (AMAX), the central UQ-flood estimates, and the central GEV estimates fitted using the 30-year record (GEV-short) and the full available record (GEV-long). The table is intended to facilitate comparisons at lower return periods, where differences among the flood-frequency curves are difficult to distinguish visually. The values represent discharge quantiles in $\text{m}^3 \text{s}^{-1}$. The Moyie River Q100 empirical estimate is reported as not available.

Table S1. Central flood-quantile estimates for return periods of 2, 5, 10, 20, 50, and 100 years for the Black River, LaHave River, and Moyie River watersheds. AMAX denotes the empirical flood-quantile estimate derived from the available long annual-maximum-flow record. UQ-flood denotes the central estimate from the proposed process-based framework developed using 30 years of observations. GEV-short denotes the central Bayesian GEV estimate based on the same 30-year observational period, and GEV-long denotes the central Bayesian GEV estimate based on the full available annual-maximum record. All discharge values are expressed in $\text{m}^3 \text{s}^{-1}$. NA indicates that an empirical Q100 estimate was not available for the Moyie River.

Watershed	Method	Q2	Q5	Q10	Q20	Q50	Q100
Black River	AMAX	128.5	155.4	176.6	194.0	223.1	230.7
	UQ-flood	128.7	163.3	183.8	202.9	226.1	242.7
	GEV-short	126.5	167.0	192.8	216.4	246.6	269.2
	GEV-long	130.1	162.1	180.3	195.8	213.4	225.1
Lahave River	AMAX	207.0	301.2	356.2	382.9	472.4	635.3
	UQ-flood	233.0	316.7	369.0	416.6	475.1	517.6
	GEV-short	225.5	348.8	460.4	602.6	867.3	1161.6
	GEV-long	215.3	301.3	363.3	427.1	517.3	591.3
Moyie River	AMAX	146.0	183.0	213.4	236.5	251.2	NA
	UQ-flood	144.3	187.2	213.4	237.7	267.0	288.1
	GEV-short	137.6	191.3	227.9	265.2	319.8	368.2
	GEV-long	146.5	189.4	212.6	231.6	252.5	265.8

S4. Diagnostic assessment of the GEV fits

The Generalized Extreme Value (GEV) distributions used as statistical benchmarks were evaluated using graphical diagnostics for each study watershed under both the short-record and long-record configurations. The assessment included quantile plots, probability plots, and return-level plots. Together, these diagnostics evaluate agreement between the fitted GEV distribution and the observed annual maxima across the central part and upper tail of the distribution.

Figures S2–S3 present the diagnostic plots for the Black River, short- and long-record fits. Figures S4–S5: Moyie River, short- and long-record fits, and Figures S6–S7: LaHave River, short- and long-record fits. The diagnostic plots show no important departures that would raise concerns about using the fitted GEV distributions as statistical benchmarks in the present analysis. Nevertheless, the GEV results remain conditional on the selected distribution, observed annual-maximum sample, and fitted parameters.

The GEV estimates are used as secondary statistical benchmarks for evaluating the central flood-quantile estimates. The diagnostic assessment does not imply that the GEV and UQ-flood uncertainty intervals represent equivalent uncertainty components or that their widths should be compared directly.

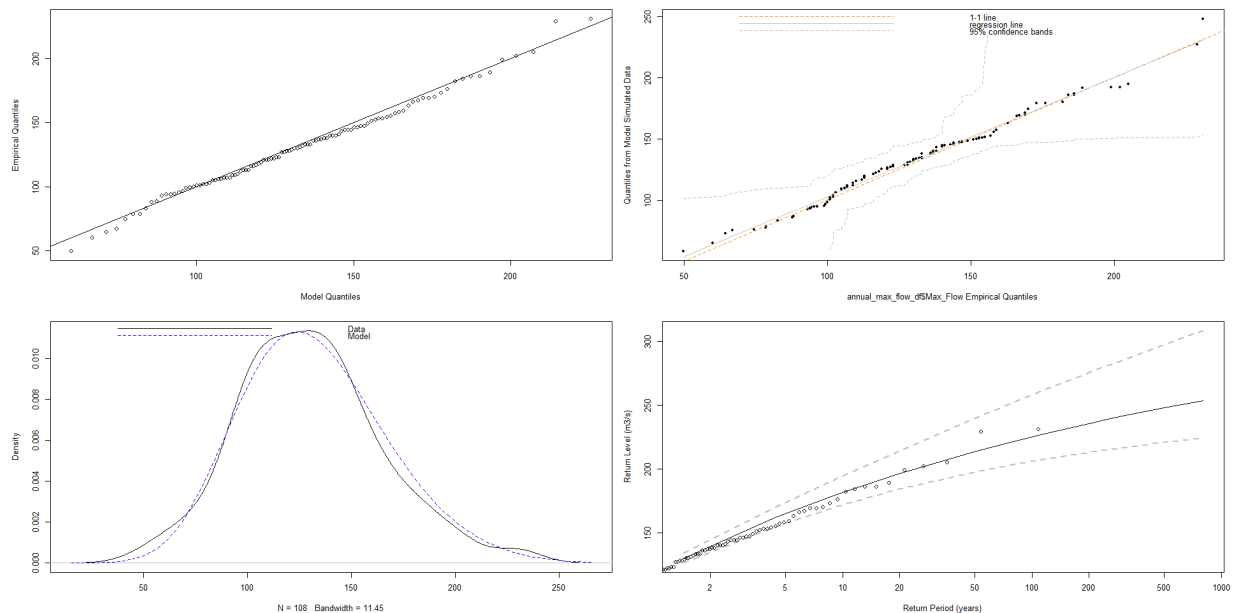


Figure S2. Goodness-of-fit diagnostics for the fitted GEV distribution for the Black River watershed under the long-record experiment, including the Q-Q plot, probability plot, density comparison, and return-level plot.

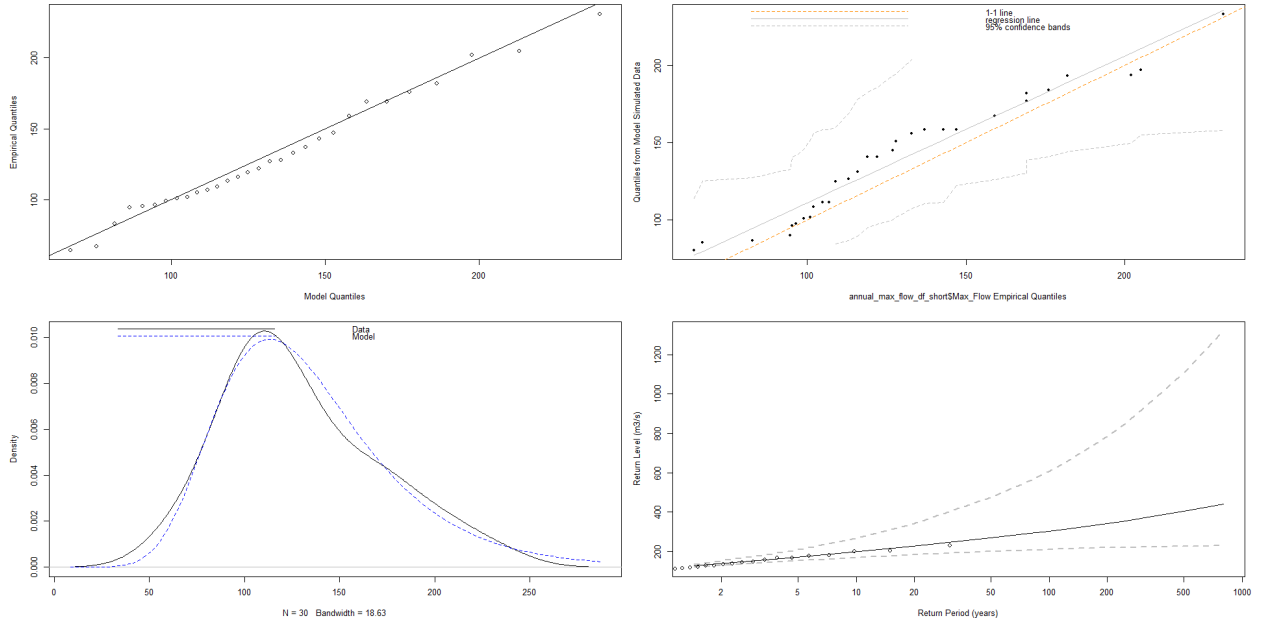


Figure S3. Goodness-of-fit diagnostics for the fitted GEV distribution for the Black River watershed under the short-record experiment, including the QQ plot, probability plot, density comparison, and return-level plot.

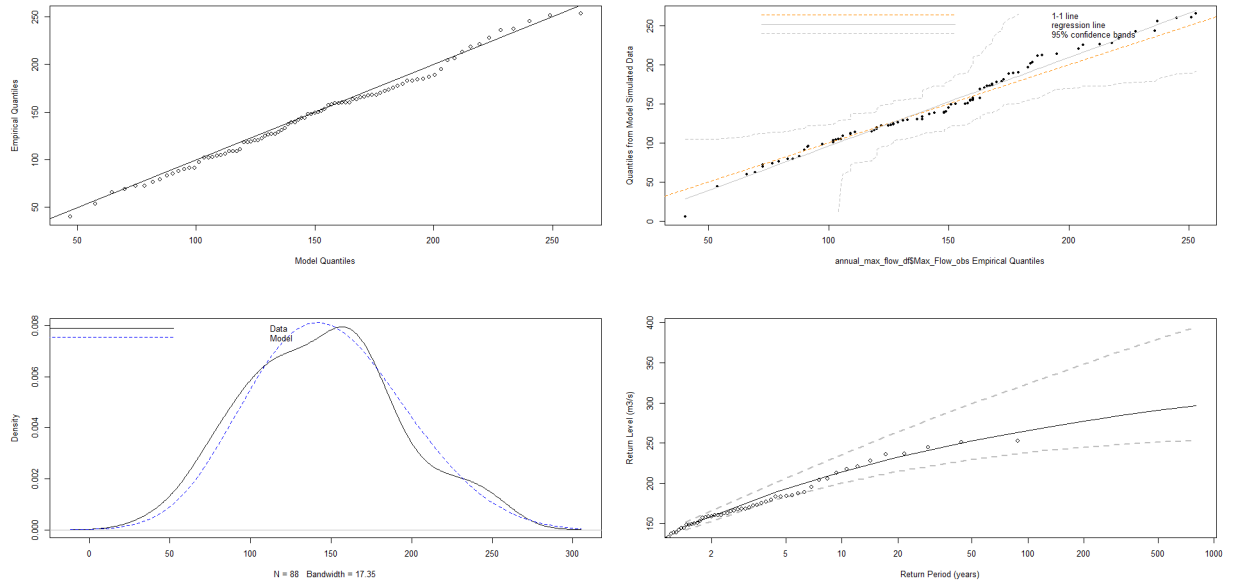


Figure S4. Goodness-of-fit diagnostics for the fitted GEV distribution for the Moyie River watershed under the long-record experiment, including the QQ plot, probability plot, density comparison, and return-level plot.

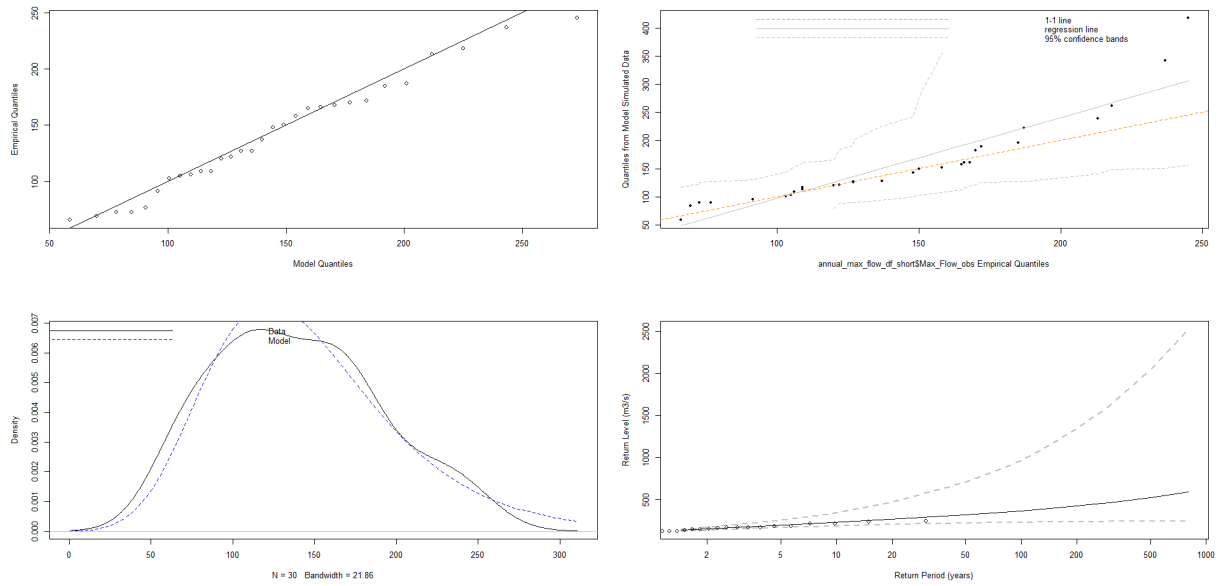


Figure S5. Goodness-of-fit diagnostics for the fitted GEV distribution for the Moyie River watershed under the short-record experiment, including the QQ plot, probability plot, density comparison, and return-level plot.

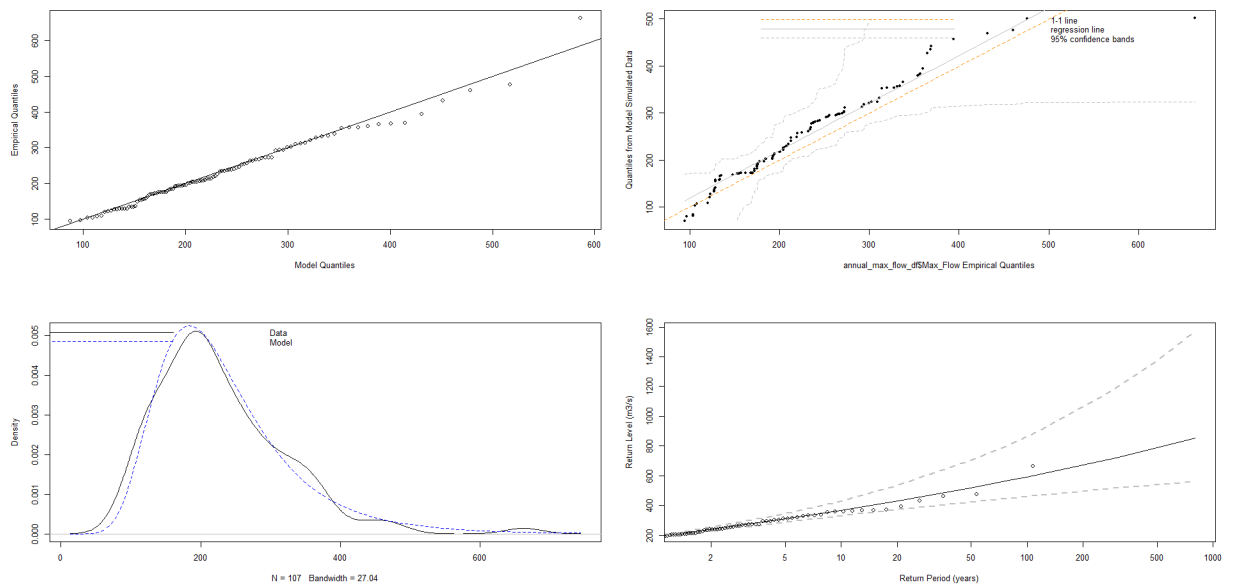


Figure S6. Goodness-of-fit diagnostics for the fitted GEV distribution for the Lahave River watershed under the long-record experiment, including the QQ plot, probability plot, density comparison, and return-level plot.

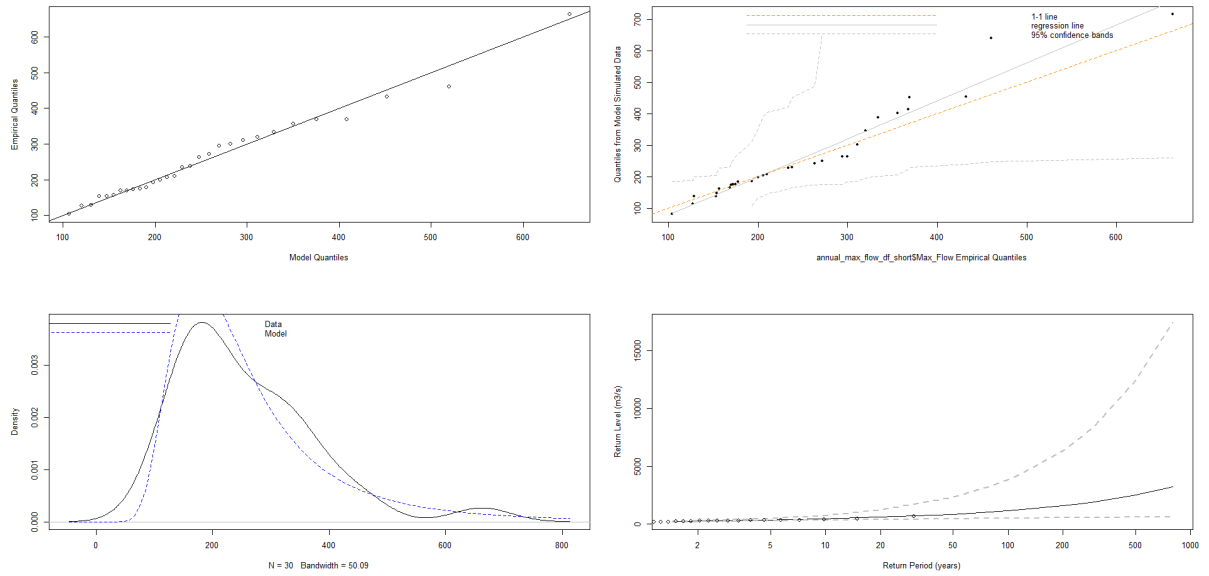


Figure S7. Goodness-of-fit diagnostics for the fitted GEV distribution for the Lahave River watershed under the short-record experiment, including the QQ plot, probability plot, density comparison, and return-level plot.

References

Cannon, A. J., Sobie, S. R., Murdock, T. Q., Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate*, 28(17), 6938–6959. <https://doi.org/10.1175/JCLI-D-14-00754.1>