ENGINEERING engrxiv ARCHIVE

GRADEX-based Curve Number calibration for HEC-HMS: an open-source R framework with bootstrap uncertainty and influence diagnostics

Statistical validation, uncertainty quantification, and automated HEC-HMS refinement for data-scarce and ungauged catchments

Mauricio Javier Victoria Niño[1]

[1]Independent Researcher, Cali, Colombia; hidratecsa@gmail.com; ORCID: 0009-0003-4328-5691

Abstract

Background and motivation. Estimating design floods in ungauged or data-scarce catchments requires hydrological models whose parameters are statistically consistent with regionalized extreme-precipitation estimates. In operational practice, the SCS Curve Number (CN) used in HEC-HMS is routinely assigned using land-use classification tables, a procedure that provides no guarantee of coherence between the frequency distribution of simulated peak discharges and that of regional extreme precipitation. The GRADEX equal-slopes criterion [1, 2] resolves this inconsistency by requiring the chord slope of the simulated discharge frequency curve (the runoff GRADEX α_Q) to match the precipitation GRADEX α_P estimated by `GRADEX-DUAL` v1.0 [3], but a standardized, reproducible computational implementation has until now been lacking.

Objective. To present `gcm_scs_gradex_calib` v1.0.0, an open-source R module that fully operationalizes the GRADEX equal-slopes CN calibration pipeline, integrating: (i) grid-search optimization with continuous CN refinement via natural cubic spline and Brent minimization, (ii) formal validation of five statistical assumptions (OLS Gumbel-linearity, Shapiro-Wilk, Durbin-Watson, LOO-CV stability), (iii) nonparametric bootstrap uncertainty quantification ($n = 2\,000$), (iv) sensitivity analysis of CN* to perturbations of α_P with local elasticity estimation, (v) LOO-based influence diagnostics per HEC-HMS simulation, and (vi) automated HEC-HMS refinement workflow when calibration error exceeds the prescribed tolerance.

Methods. For each CN on the user-defined search grid the module computes $\alpha_Q = [Q(T{=}100) - Q(T{=}10)] \,/\, \Delta y$ and minimizes $E(\text{CN}) = |\alpha_Q - \alpha_P| \,/\, \alpha_P$. The discrete optimum is refined to CN$^*_{\text{cont}}$ via natural cubic spline interpolation [7] and Brent's method [8]. Statistical validation comprises OLS regression of ΔQ vs. the Gumbel reduced variate ($R^2 \geq 0.95$), Shapiro-Wilk normality and Durbin-Watson autocorrelation tests, and LOO-CV with empirical 95 % confidence intervals [9]. Bootstrap uncertainty ($n = 2\,000$ replicates) and a structured α_P perturbation sweep ($\pm 30\,\%$, step 5 %) complete the statistical analysis.

Results. Applied to a ten-simulation HEC-HMS dataset (CN $\in [50, 95]$, step $= 5$; $\alpha_P = 20.397$ mm from `GRADEX-DUAL` v1.0), the module yields CN$^*_{\text{cont}}$ $= 71.289$ (discrete: 70.0; $E = 11.33\,\%$). The error exceeds the 5 % tolerance because the CN grid step is too coarse to resolve the optimum; the

module automatically generates a HEC-HMS refinement work list. The calibration is confirmed stable (LOO-CV SD $= 1.581$; 95 % CI $= [70.0,\ 73.9]$), and bootstrap gives a 95 % CI of $[65.0,\ 75.0]$ ($n = 2\,000$; skewness $= -0.243$). The CN* sensitivity range over $\pm 30\,\%$ GRADEX uncertainty is $[65.2,\ 76.5]$. CN $= 70$ is identified as the sole highly influential simulation (influence index $= 1.0$, ΔCN $= +5.0$).
Conclusions. `gcm_scs_gradex_calib` v1.0.0 provides a fully automated, reproducible and auditable pipeline for GRADEX-based CN calibration in HEC-HMS. The module is particularly valuable in data-scarce and ungauged catchments where regionalized extreme-precipitation GRADEX is available but streamflow records are absent or insufficient for direct flood frequency analysis.

Keywords: GRADEX; Curve Number calibration; HEC-HMS; SCS/CN; flood frequency analysis; equal-slopes criterion; extreme hydrology; bootstrap uncertainty; influence diagnostics; open-source R software.
Software availability: `gcm_scs_gradex_calib` v1.0.0 is open-source under a CC BY 4.0 license. The R script is available at https://github.com/MauricioVictoriaN/GCM-SCS-GRADEX-CALIB. The companion module `GRADEX-DUAL` v1.0 [3] is published at doi:10.31224/6945 and its code is available at https://github.com/MauricioVictoriaN/GRADEX-DUAL.

1 Introduction

The estimation of design floods in ungauged or data-scarce catchments is one of the central challenges of applied hydrology [16]. Semi-distributed models such as HEC-HMS [4] require careful parameterization of multiple components, including concentration time (T_c), initial abstraction (I_a), and the SCS Curve Number (CN), a dimensionless index that integrates the effects of soil type, land use and antecedent moisture condition on the generation of direct runoff [5]. In operational practice, CN is typically assigned using tabulated land-use classifications, a procedure that provides no guarantee of statistical coherence between the simulated peak-discharge frequency distribution and the regional extreme-precipitation frequency distribution [2].

The GRADEX method [1] provides a theoretically grounded framework linking extreme precipitation to flood frequency through the asymptotic behavior of the Gumbel (EV1) distribution. At high return periods, the annual-maximum precipitation approaches the Gumbel asymptote, and the GRADEX parameter α_P — the scale parameter of that asymptote — characterizes the rate of growth of extreme precipitation with return period [1]. The equal-slopes criterion [2] extends this result to simulated streamflows: when the soil is near saturation (CN $\to$ 100), the SCS/CN transformation introduces only a constant additive shift, so the annual-maximum discharge distribution inherits the same Gumbel slope as precipitation. This restriction uniquely determines the CN value — denoted CN* — that makes HEC-HMS statistically consistent with the regional extreme-precipitation estimate, without requiring streamflow records at the target site.

The precipitation GRADEX α_P is supplied by `GRADEX-DUAL` v1.0 [3], a companion open-source R module published at doi:10.31224/6945 that estimates α_P from point-rainfall or regional L-moment analysis [6], together with a $\pm 30\,\%$ sensitivity range.

Despite the theoretical elegance and operational appeal of the equal-slopes calibration, existing applications [2] rely on manual or semi-manual procedures that are not easily auditable, reproducible, or extendable to include modern statistical diagnostics such as bootstrap uncertainty quantification or influence diagnostics. This gap is particularly limiting in regulatory contexts where the calibration must be fully documented, or in comparative studies requiring replication across multiple catchments.

This paper presents `gcm_scs_gradex_calib` v1.0.0, an open-source R module that operationalizes the complete GRADEX equal-slopes calibration pipeline as a single reproducible script. The specific objectives are listed in Section 1.1.

1.1 Objectives

1. Implement the GRADEX equal-slopes criterion [1, 2] as a reproducible, documented R module integrating directly with `GRADEX-DUAL` v1.0 [3].

2. Develop a continuous CN* estimator via natural cubic spline interpolation and Brent minimization, reducing discretization error from $O(\pm\Delta\text{CN}/2)$ to $O(10^{-3})$ CN units.

3. Validate the five statistical assumptions of the GRADEX method through OLS Gumbel-linearity regression, Shapiro-Wilk normality and Durbin-Watson autocorrelation tests, and LOO-CV calibration stability.

4. Quantify CN* sampling uncertainty via non-parametric bootstrap ($n = 2\,000$) and characterize its sensitivity to α_P perturbations through a structured sweep.

5. Identify influential HEC-HMS simulations through a LOO influence index and provide automated guidance for simulation refinement when calibration error exceeds the prescribed tolerance.

6. Provide publication-quality diagnostic outputs: multi-panel figures, a structured six-sheet Excel workbook, and a plain-text technical report with full metadata traceability.

2 Theoretical Framework

2.1 The GRADEX parameter

For large return periods, the annual-maximum precipitation distribution converges asymptotically to the Gumbel (EV1) distribution, making the quantile function asymptotically linear in the Gumbel reduced variate $y(T)$ [1]:

$$Q_P(T) \sim u_P + \alpha_P \cdot y(T), \qquad y(T) = -\ln\big(-\ln(1 - T^{-1})\big), \tag{1}$$

where u_P is the Gumbel location parameter (mm), α_P is the GRADEX or scale parameter (mm), and $y(T)$ is the Gumbel reduced variate. The GRADEX is independent of storm duration for a given climatic region [1] and is estimated by `GRADEX-DUAL` v1.0 [3].

2.2 The equal-slopes criterion

Under the SCS/CN runoff transformation, as the potential maximum retention $S \to 0$ (equivalently CN $\to$ 100), the effective rainfall equals gross rainfall minus a constant, so the annual-maximum discharge distribution inherits the Gumbel slope of precipitation [2]:

$$Q_Q(T) \sim u_Q + \alpha_Q \cdot y(T), \qquad \alpha_Q = \alpha_P. \tag{2}$$

The calibration criterion follows from Eq. (2):

$$\text{CN}^* = \arg\min_{\text{CN}} \big|\alpha_Q(\text{CN}) - \alpha_P\big|. \tag{3}$$

2.3 Runoff GRADEX estimator

For each CN on the search grid, HEC-HMS provides $Q(T{=}10)$ and $Q(T{=}100)$ in mm (depth over catchment area). Following [2], the $T = 10$ yr and $T = 100$ yr quantiles define a chord spanning the operational design range; the resulting $\Delta y = 2.34978$ is sufficiently large to minimize the influence of estimation uncertainty on α_Q. The runoff GRADEX is estimated as the chord slope on the Gumbel probability plot:

$$\alpha_Q = \frac{Q(T = 100) - Q(T = 10)}{y(100) - y(10)}, \tag{4}$$

where $y(10) = 2.25037$, $y(100) = 4.60015$ and $\Delta y = 4.60015 - 2.25037 = 2.34978$ (dimensionless).[1] Since Q is expressed in mm and Δy is dimensionless, both α_Q and α_P carry units of mm per unit of Gumbel reduced variate (mm $[\Delta y]^{-1}$), and their ratio in Eq. (5) is therefore dimensionless.

2.4 Calibration error and tolerance

The relative calibration error is:

$$E(\text{CN}) = \frac{|\alpha_Q(\text{CN}) - \alpha_P|}{\alpha_P} \times 100 \quad [\%]. \tag{5}$$

[1] The Gumbel reduced variate is $y(T) = -\ln[-\ln(1 - 1/T)]$; $y(10) = -\ln(-\ln 0.9) = 2.25037$, $y(100) = -\ln(-\ln 0.99) = 4.60015$, $\Delta y = 4.60015 - 2.25037 = 2.34978$. Values rounded to five decimal places.

Calibration is declared successful when $E(\text{CN}^*) \leq 5\,\%$ (Eq. (3); [2]).

2.5 Continuous CN refinement

A natural cubic spline [7] is fitted to the three grid points surrounding the discrete minimum, and Brent's method [8] locates the continuous global minimum via `stats::optimize()` in R:

$$\text{CN}^*_{\text{cont}} = \arg \min_{[\text{CN}_{i-1},\, \text{CN}_{i+1}]} f_{\text{spline}}(\text{CN}). \tag{6}$$

Three robustness guards are applied to Eq. (6): (G1) $\text{CN}^*_{\text{cont}}$ is clamped to $[1, 100]$ if the spline root falls outside physical bounds; (G2) a warning is issued if $\text{CN}^*_{\text{cont}}$ deviates from the discrete optimum by more than $2 \times \Delta$CN; (G3) the spline step is skipped and a warning issued if the range of α_Q across the grid is less than 0.01 mm (flat error curve).

2.6 Statistical assumptions of the method

Five assumptions underpin the GRADEX equal-slopes calibration. Table 1 outlines each assumption alongside its respective verification mechanism.

Table 1: Statistical assumptions of the GRADEX equal-slopes criterion and their verification in `gcm_scs_gradex_calib` v1.0.0.

Ass.	Statement	Reference / Verification
A1	Saturated soil: exact only as CN $\to$ 100; conservative for CN$^* < 80$ (author's interpretation based on the asymptotic derivation in [2]).	[2]
A2	Simulated discharges follow the Gumbel (EV1) distribution.	OLS regression, $R^2 \geq 0.95$ [10]
A3	Independence of (Q_{10}, Q_{100}) pairs with respect to CN ordering.	Durbin-Watson test [11, 12]
A4	Stationarity of HEC-HMS parameters over $T \in [10, 100]$ yr.	Documented verification
A5	Strict monotonicity of Q(CN).	`.validate_hec_simulations()` [5]

3 Statistical and Numerical Methods

3.1 Gumbel linearity regression

An OLS regression is fitted to detect departure from Gumbel linearity:

$$\Delta Q = a + b \cdot y_{\text{theo}} + \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0,\, \sigma^2), \tag{7}$$

where $\Delta Q = Q_{100} - Q_{10}$ (mm) and y_{theo} maps the normalized CN rank to the Gumbel scale $[y_{10},\, y_{100}]$. Linearity is accepted when $R^2 \geq 0.95$ in Eq. (7) [10]. An informational Pearson r between the normalized CN and ΔQ is reported but does not enter the linearity decision.

3.2 Normality and autocorrelation tests

- **Shapiro-Wilk** [13]: tests normality of OLS residuals (H_0: normality; rejected at $p < 0.05$).
- **Durbin-Watson** [11, 12]: tests first-order autocorrelation of OLS residuals (H_0: no autocorrelation; expected $p > 0.05$). For residuals ordered by the monotone CN covariate, the DW test detects system-

atic departures from OLS linearity rather than serial dependence [14]; a significant result therefore flags lack of linear fit (i.e., a nonlinearity signal) rather than temporal autocorrelation. The residual plot (Fig. 1) is the primary diagnostic; the DW result serves as a secondary, quantitative flag.

3.3 Leave-One-Out cross-validation

CN^* stability is assessed by iteratively excluding one HEC-HMS simulation and re-calibrating on the remaining $n-1$ points [9]. The stability criterion is $\text{SD}_{\text{LOO}} < 2.0$ CN units; the 95 % CI is computed from empirical percentiles of the LOO distribution.

3.4 Bootstrap uncertainty quantification

Non-parametric bootstrap [9] with $n = 2\,000$ replicates (configurable; minimum 100) characterizes the sampling distribution of CN^*. In each replicate, the HEC-HMS simulation table is resampled with replacement and $\text{CN}^*_{\text{cont}}$ re-estimated. Summary statistics (mean, median, SD, skewness, empirical 95 % CI) and the full bootstrap vector are reported. With small simulation sets (fewer than 8 CN grid points), resampling may produce replicates with duplicate CN values and flat error curves (zero gradient in $E(\text{CN})$); such replicates are excluded and counted in `na_count`. The case study uses 10 CN grid points, yielding $2\,000/2\,000$ valid replicates because with $n = 10$ the probability of drawing a flat configuration is negligible ($< 10^{-6}$). A minimum of 8 CN grid points is recommended for reliable bootstrap inference.

3.5 Sensitivity of CN^* to α_P

For each percentage perturbation $\delta \in \Delta$ (default: $[-30, -25, \ldots, +30]\,\%$, step 5 %), the perturbed target $\alpha_P^{\delta} = \alpha_P(1 + \delta/100)$ is formed and $\text{CN}^*_{\text{cont}}$ re-estimated. The local elasticity is computed via a centered finite-difference approximation:

$$\varepsilon_i = \frac{\Delta \text{CN}^*}{\Delta \alpha_P} \times \frac{\alpha_{P,i}}{\text{CN}^*_i}. \tag{8}$$

The elasticity is a local approximation whose reliability improves with denser perturbation grids; the default step of 5 % represents a practical compromise between resolution and computational cost.

3.6 LOO influence diagnostics

An influence index is computed for each HEC-HMS simulation by comparing the LOO CN^* against the full-dataset discrete CN^* [15]:

$$I_j = \frac{|\text{CN}^*_{\text{LOO}(-j)} - \text{CN}^*_{\text{disc}}|}{\max_k |\text{CN}^*_{\text{LOO}(-k)} - \text{CN}^*_{\text{disc}}|}. \tag{9}$$

Simulations with $I_j > 0.5$ (Eq. (9)) are flagged as highly influential.

3.7 Automated HEC-HMS refinement

When $E > 5\,\%$, the tolerance band on α_Q is:

$$\alpha_Q^{\pm} = \alpha_P \left(1 \pm \frac{\text{tol}}{100}\right), \tag{10}$$

and the corresponding CN interval is determined by inverting Eq. (10) via spline interpolation. A recommended simulation range is constructed by expanding the strict interval by two grid steps on each side. A CSV template listing existing and new required HEC-HMS simulations is exported automatically.

4 Software Architecture

4.1 Design principles

The module follows a pure-function architecture with no mutable global state. All parameters — including all input/output file paths and names — are declared as configurable constants in Section 1 of the script; no hard-coded paths appear elsewhere in the code. The reproducibility seed (`RANDOM_SEED`) is applied via `set.seed()` immediately before the bootstrap loop; Brent optimization and spline fitting are deterministic and require no seed. Output files (including `hec_refinement_template.csv`) are overwritten on each run; users should archive previous outputs before re-running. The implementation requires R $\geq$ 4.1.0 and three external packages: `readxl` [18], `writexl` [19] and `ggplot2` [20], eliminating the Java runtime dependency of legacy Excel packages. A step-consistency hard stop in `.validate_hec_simulations()` halts execution if the CN step inferred from the loaded simulation file differs from the value declared in `CN_STEP`.

4.2 Module structure

The module is organized in 13 logical sections. Table 2 lists all public functions and Table 3 lists input/output files.

Table 2: Public functions of `gcm_scs_gradex_calib` v1.0.0.

Function	Purpose
`gcm_gumbel_variate(T)`	Gumbel reduced variate $y(T)$ (Eq. (1))
`gcm_runoff_gradex(q10, q100, dy)`	Chord-slope runoff GRADEX α_Q (Eq. (4))
`gcm_calib_relative_error(aq, ap)`	Relative calibration error E (Eq. (5))
`gcm_read_gradex_dual(path)`	Read α_P from GRADEX-DUAL v1.0 workbook
`gcm_generate_hec_template(...)`	Generate a blank CSV template for HEC-HMS simulation inputs
`gcm_example_hec_data(...)`	Generate a synthetic HEC-HMS dataset for pipeline verification; invoked internally by `gcm_load_hec_simulations(use_example = TRUE)`
`gcm_load_hec_simulations(path,...)`	Load, validate and step-check HEC-HMS file
`gcm_calibrate_cn(sim_df, alpha_p)`	Grid search + spline/Brent for CN* with guards G1–G3
`gcm_validate_gumbel_assumptions()`	OLS R^2, Shapiro-Wilk, Durbin-Watson, Pearson r
`gcm_loo_cn_stability()`	LOO-CV stability of CN*
`gcm_bootstrap_cn_uncertainty()`	Bootstrap CI ($n = 2\,000$)
`gcm_cn_sensitivity_to_gradex()`	CN* sensitivity sweep and local elasticity
`gcm_identify_influential_simulations()`	LOO influence index per simulation
`gcm_recommend_hec_refinement()`	Refinement range and HEC-HMS work list
`gcm_plot_gumbel_linearity()`	OLS diagnostic figure
`gcm_plot_cn_calibration()`	Five-panel calibration figure
`gcm_plot_cn_sensitivity()`	Sensitivity and elasticity figure
`gcm_export_results_xlsx()`	Six-sheet Excel workbook
`gcm_generate_technical_report()`	Plain-text technical report

Table 3: Input and output files of `gcm_scs_gradex_calib` v1.0.0.

File	Dir.	Description
`resultados_gradex_completos.xlsx`	Input	GRADEX-DUAL v1.0 workbook
`results_hec_hms.csv/.xlsx`	Input	HEC-HMS simulations (Q_{10}, Q_{100})
`diagnostic_gumbel_linearity.png`	Output	OLS Gumbel linearity diagnostic
`cn_calibration_gradex.png`	Output	Five-panel calibration figure
`cn_sensitivity_gradex.png`	Output	Sensitivity and elasticity figure
`cn_calibration_gradex.xlsx`	Output	Six-sheet Excel workbook
`calibration_results.csv`	Output	Full calibration table
`influential_simulations.csv`	Output	Influence index per simulation
`hec_refinement_template.csv`	Output	HEC-HMS work list (conditional)
`calibration_report.txt`	Output	Plain-text technical report

The six-sheet Excel workbook produced by `gcm_export_results_xlsx()` contains: *Summary* — 34-row metadata block with input parameters, calibration metrics and assumption-validation status; *Results* — full $(CN,\ \alpha_Q,\ E)$ table; *LOO_CV* — per-fold CN* estimates; *Bootstrap* — bootstrap replicate statistics and distribution summary; *Influence_Analysis* — LOO influence index per simulation; and *Refinement* — recommended HEC-HMS simulation range (omitted when calibration is successful).

4.3 User-configurable parameters

Table 4: User-configurable parameters of `gcm_scs_gradex_calib` v1.0.0.

Parameter	Default	Justification
`CN_MIN`	50	Recommended lower bound [5]
`CN_MAX`	95	Recommended upper bound
`CN_STEP`	5.0	Grid step (must match HEC-HMS runs)
`CALIB_TOL_PCT`	5.0 %	Engineering convention [2]; adjust to project requirements
`LINEARITY_R2_MIN`	0.95	Gumbel linearity threshold [10]
`ALPHA_SIG`	0.05	Statistical significance level [21]
`n_boot`	2 000	Bootstrap replicates (function argument) [9]
`delta_pct`	±30 %, step 5	Perturbation sweep range (function argument)
`RANDOM_SEED`	2024	Reproducibility seed for bootstrap

5 Outputs and Diagnostic Graphics

The module generates three PNG figures at $300\,\mathrm{DPI}$.

5.1 Gumbel linearity diagnostic

Fig. 1 evaluates Assumption A2 by plotting $\Delta Q = Q_{100} - Q_{10}$ against the theoretical Gumbel variate y_{theo} mapped to the normalized CN rank. The OLS fit is shown with a 95 % confidence envelope. Points should lie close to a line with positive slope; systematic curvature indicates inappropriate HEC-HMS parameterization of concentration time (T_c) or initial abstraction (I_a).

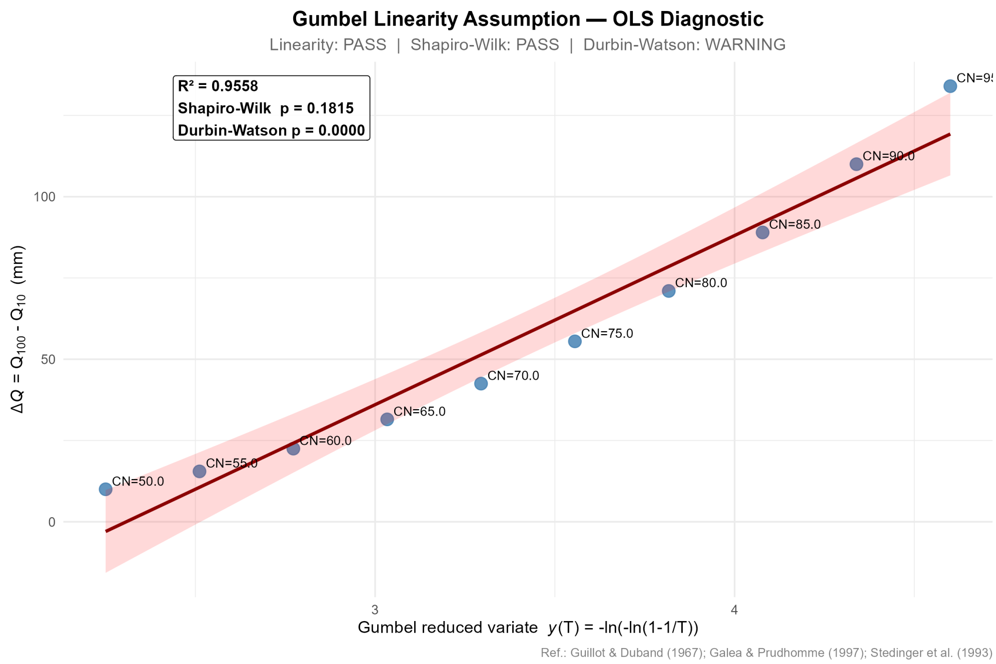

Figure 1: Gumbel linearity diagnostic. OLS regression of ΔQ vs. y_{theo} with 95 % confidence envelope. Labels show CN values. Case study: $R^2 = 0.9558$ (PASS; threshold ≥ 0.95), Shapiro-Wilk $p = 0.1815$ (PASS), Durbin-Watson $p < 0.001$ (WARNING).

5.2 Five-panel calibration figure

Fig. 2 synthesizes the optimization and uncertainty analysis in five panels: (1) $\alpha_Q(\text{CN})$ versus target α_P with tolerance and GRADEX sensitivity bands; (2) relative error $E(\text{CN})$; (3) CN uncertainty by GRADEX scenario; (4) LOO-CV stability; (5) bootstrap distribution of CN*.

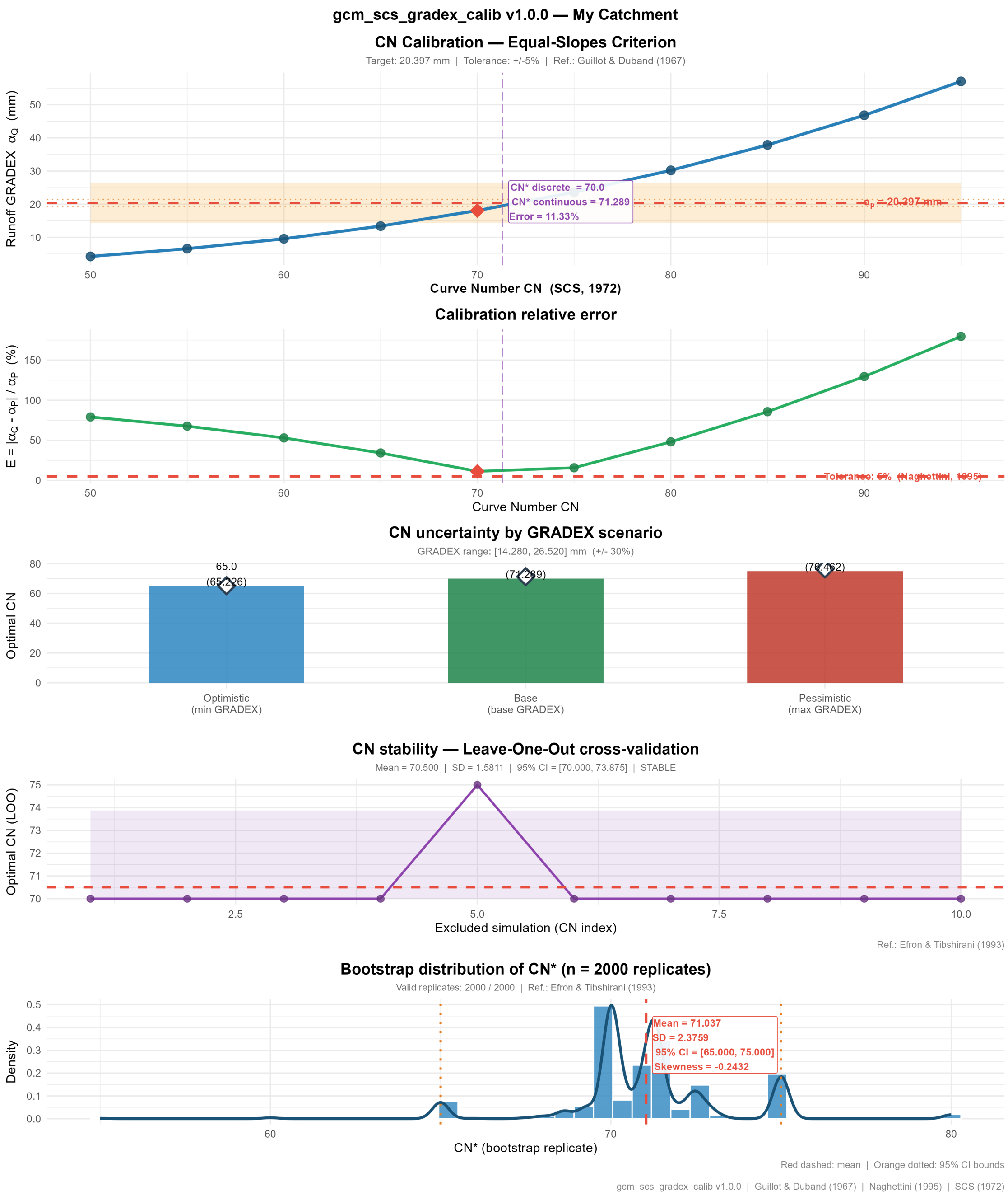

Figure 2: Five-panel GRADEX-SCS/CN calibration dashboard ($\alpha_P = 20.397$ mm, $\text{CN}^*_{\text{cont}} = 71.289$, $E = 11.33\,\%$, LOO-CV SD $= 1.581$, bootstrap 95 % CI $= [65.0,\ 75.0]$, $n = 2\,000$). Panels: (1) $\alpha_Q(\text{CN})$ vs. target α_P with $\pm 5\,\%$ tolerance and $\pm 30\,\%$ GRADEX sensitivity bands; (2) relative error $E(\text{CN})$; (3) CN* by GRADEX scenario; (4) LOO-CV stability; (5) bootstrap distribution of CN*.

5.3 Sensitivity and elasticity figure

Fig. 3 presents the two-panel sensitivity analysis. The upper panel plots $\text{CN}^*_{\text{cont}}$ versus perturbed α_P, distinguishing successful calibrations ($E \leq 5\,\%$, green) from failures (red). The lower panel shows the local elasticity (Eq. (8)).

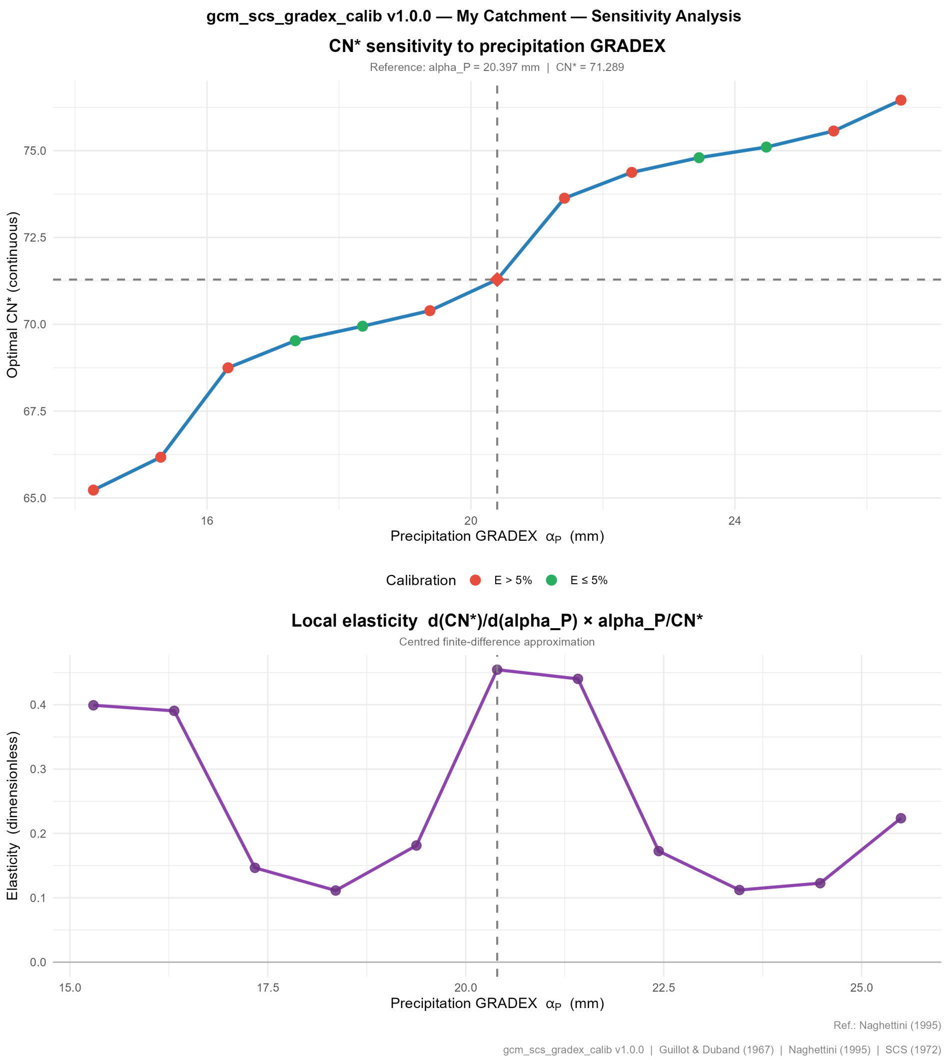

Figure 3: CN* sensitivity to precipitation GRADEX α_P. Upper panel: $\text{CN}^*_{\text{cont}}$ vs. perturbed α_P; green: $E \leq 5\,\%$; red: $E > 5\,\%$. Lower panel: local elasticity $\varepsilon = (\text{d CN}^*/\text{d}\alpha_P) \times (\alpha_P/\text{CN}^*)$.

6 Case Study

6.1 Input data

The module is demonstrated with a ten-simulation HEC-HMS dataset spanning CN $\in [50, 95]$ at a grid step of 5 CN units. The precipitation GRADEX was estimated by `GRADEX-DUAL` v1.0 [3], yielding $\alpha_P = 20.397$ mm (source: `HW_Quality_Diag` sheet; sensitivity range $= [14.280,\ 26.520]$ mm, $\pm 30\,\%$). Table 5 summarizes the simulation inputs and derived α_Q.

Note on the case study design. The dataset used in this demonstration was deliberately configured with a coarse CN grid step (ΔCN $= 5$) to exercise the full range of diagnostic and corrective capabilities of the module — including the calibration-error warning ($E > 5\,\%$), the HEC-HMS refinement workflow, and the influence diagnostics. In operational use, practitioners would typically employ a finer grid step (ΔCN ≤ 2) that resolves the optimum within tolerance. The "POOR" quality rating reported by the module is therefore an expected and intended outcome of this controlled test configuration, not a limitation of the method itself. Additionally, the module provides the function `gcm_example_hec_data()` (invoked via `gcm_load_hec_simulations(use_example = TRUE)`), which generates a fully synthetic HEC-HMS dataset for pipeline verification without requiring any external simulation files. This self-contained testing mode enables users to validate the installation and inspect all diagnostic outputs before substituting their own HEC-HMS results.

Table 5: HEC-HMS simulation dataset and derived runoff GRADEX values. $\Delta y = 2.34978$.

CN	Q_{10} (mm)	Q_{100} (mm)	α_Q (mm)	E (%)
50	8.0	18.0	4.256	79.14
55	12.0	27.5	6.596	67.66
60	17.5	40.0	9.575	53.06
65	24.5	56.0	13.405	34.28
70	33.0	75.5	18.087	11.33
75	43.5	99.0	23.619	15.80
80	56.0	127.0	30.216	48.14
85	71.0	160.0	37.876	85.69
90	88.5	198.5	46.813	129.5
95	109.0	243.0	57.027	179.6

6.2 Calibration results

Table 6 summarizes calibration results for the three GRADEX scenarios. The base-scenario error of 11.33 % exceeds the 5 % tolerance because the discrete grid step of 5 CN units is too coarse to resolve the optimum, which lies between CN $= 70$ and CN $= 75$. Importantly, the continuous estimator $\text{CN}^*_{\text{cont}} = 71.289$ does not reduce the error below 11.33 %: spline interpolation refines the location of the minimum within the existing grid, but cannot recover information absent from the simulation set. The grid itself must be made finer, as recommended by the automated refinement workflow (Section 6.6).

Table 6: Calibration results by GRADEX scenario. α_P from `GRADEX-DUAL` v1.0 [3].

Scenario	α_P **(mm)**	$\text{CN}^*_{\text{disc}}$	$\text{CN}^*_{\text{cont}}$	α_Q **(mm)**	E **(%)**
Optimistic	14.280	65.0	65.226	13.405	6.12
Base	20.397	70.0	71.289	18.087	11.33
Pessimistic	26.520	75.0	76.462	23.619	10.94

Table 7: Step-by-step calculation of α_Q for each CN in the search grid. $\Delta y = 2.34978$ (dimensionless).

CN	$\mathbf{Q_{10}}$ (mm)	$\mathbf{Q_{100}}$ (mm)	$\mathbf{\Delta Q}$ (mm)	$\alpha_\mathbf{Q} = \mathbf{\Delta Q/\Delta y}$ (mm)	$\mathbf{E}(\%)$
50	8.0	18.0	10.0	4.256	79.14
55	12.0	27.5	15.5	6.596	67.66
60	17.5	40.0	22.5	9.575	53.06
65	24.5	56.0	31.5	13.406	34.28
70	33.0	75.5	42.5	18.087	11.33
75	43.5	99.0	55.5	23.619	15.80
80	56.0	127.0	71.0	30.216	48.14
85	71.0	160.0	89.0	37.876	85.69
90	88.5	198.5	110.0	46.813	129.5
95	109.0	243.0	134.0	57.027	179.6

6.3 Statistical assumption validation

Assumption A2 (Gumbel linearity) is verified: $R^2 = 0.9558 \geq 0.95$ (F-test $p < 10^{-4}$). Residual normality (Shapiro-Wilk test) is verified: $p = 0.1815 > 0.05$. Assumption A3 (residual independence, Durbin-Watson) triggers a warning ($p < 0.001$), flagging a nonlinearity signal in the OLS residuals consistent with the curvature of the CN-to-ΔQ mapping (see Section 6.8). The informational Pearson $r = 0.9777$ between normalized CN and ΔQ is not used in the linearity decision.

6.4 LOO-CV stability and bootstrap uncertainty

LOO-CV: mean $\text{CN}^*_{\text{LOO}} = 70.500$; SD $= 1.581$ (**STABLE**, threshold < 2.0); empirical 95 % CI $= [70.0,\ 73.9]$; 10/10 valid replicates. Bootstrap ($n = 2\,000$, seed $= 42$): mean $= 71.037$; median $= 71.177$; SD $= 2.376$; skewness $= -0.243$; 95 % CI $= [65.0,\ 75.0]$; $2\,000/2\,000$ valid replicates.

6.5 Influence diagnostics

Table 8 reports the LOO influence index. CN $= 70$ is the sole highly influential simulation ($I = 1.0$, ΔCN $= +5.0$): excluding it forces the optimum to CN $= 75$. All remaining simulations have $I = 0$.

Table 8: LOO influence diagnostics. $\text{CN}^*_{\text{disc}} = 70.0$ (full dataset). Simulations with $I > 0.5$ are highly influential.

CN	$\text{CN}^*_{\text{LOO}(-j)}$	ΔCN	Index I	Highly influential
70	75	5.000	1.000	Yes
50	70	0.000	0.000	No
55	70	0.000	0.000	No
60	70	0.000	0.000	No
65	70	0.000	0.000	No
75	70	0.000	0.000	No
80	70	0.000	0.000	No
85	70	0.000	0.000	No
90	70	0.000	0.000	No
95	70	0.000	0.000	No

6.6 HEC-HMS refinement recommendation

The tolerance band $\alpha_Q \in [19.377,\ 21.417]$ mm corresponds to a strict CN interval of $\approx [71.2,\ 73.0]$, entirely between the CN $= 70$ and CN $= 75$ grid points. The recommended simulation range is CN $\in [60,\ 85]$ at step 5, reusing the six existing simulations that fall within this range (a subset of the ten in Table 5); no new HEC-HMS runs are required. The refinement template (`hec_refinement_template.csv`) is exported automatically.

6.7 Sensitivity analysis

Over the $\pm 30\,\%$ GRADEX band (13 perturbation levels), 4 of 13 calibrations satisfy $E \leq 5\,\%$. The CN* range across the sweep is $[65.2,\ 76.5]$ (11.3 CN units), directly usable as the HEC-HMS sensitivity range [2]. The local elasticity at the base case is $\varepsilon \approx 0.45$: a 10 % perturbation of α_P produces approximately a 4.5 % change in CN*.

6.8 Limitations of the case study

The Durbin-Watson warning reflects the inherent nonlinearity of the CN-to-ΔQ mapping that the OLS linearization approximates; it does not invalidate the calibration but indicates that assumption A2 is met only marginally at the boundary of its validity range. The primary limitation is the coarse CN grid step of 5 units, which prevents achieving $E \leq 5\,\%$ without additional HEC-HMS simulations at intermediate CN values.

7 Results Interpretation Guide

7.1 Calibration quality matrix

Table 9: Calibration quality matrix: recommended actions by indicator combination.

Indicator combination	Status	Recommended action
$E \leq 5\,\%$, $R^2 \geq 0.95$, $\text{SD}_{\text{LOO}} < 2$	EXCELLENT	Use $\text{CN}^*_{\text{cont}}$ directly in HEC-HMS; conduct sensitivity analysis over bootstrap 95 % CI.
$E \leq 5\,\%$, $R^2 \geq 0.95$, $\text{SD}_{\text{LOO}} \geq 2$	ACCEPTABLE	Inspect LOO deviations and influential simulations; consider expanding the CN search range.
$E \leq 5\,\%$, $R^2 < 0.95$	LIMITED	Review HEC-HMS parameterization (T_c, I_a); consider alternative extreme-value distributions (GEV, PE3).
$E > 5\,\%$	NEEDS REFINEMENT	Reduce CN grid step (see `hec_refinement_template.csv`); verify α_P from GRADEX-DUAL v1.0.

7.2 Discrete vs. continuous CN*

$\text{CN}^*_{\text{cont}}$ should always be preferred for HEC-HMS entry, as it minimizes discretization error. When Guard G2 is triggered ($\text{CN}^*_{\text{cont}}$ deviates from $\text{CN}^*_{\text{disc}}$ by more than $2 \times \Delta$CN), spline extrapolation is unreliable and the discrete value should be used instead.

8 Limitations and Scope

- `gcm_scs_gradex_calib` calibrates CN only. Prior to applying the module, all other HEC-HMS parameters (T_c, I_a, routing coefficients) must be fixed by independent means (regionalization, tables, or separate calibration). The module assumes these parameters are correct and stable.
- The equal-slopes asymptote is exact only at full saturation (CN $\rightarrow$ 100); for lower CN (as in the case study, where CN$^* = 71.3 < 80$) the transformation introduces an additive non-constant shift that inflates α_Q relative to α_P, and the method may overestimate extreme discharges [2]. The analyst should report CN* relative to the 80-unit threshold as a precautionary indicator.
- CN is calibrated as a lumped catchment parameter; spatial heterogeneity may require sub-catchment calibration.
- The target α_P must come from a precipitation network representative of the study catchment; uncertainty from `GRADEX-DUAL` v1.0 should be propagated and reported alongside CN*.
- Climatic non-stationarity and land-use change during the calibration period are not addressed; for non-stationary contexts, GAMLSS-based approaches [17] should be considered.
- LOO-CV measures internal calibration stability, not predictive performance on independent events; external validation with observed flood records remains essential where available.
- The module treats HEC-HMS as a deterministic black box that provides Q_{10} and Q_{100} for each CN. Uncertainties arising from the design hyetograph, the transformation method (e.g., SCS unit hydrograph, kinematic wave), the time step, and the numerical solver are *not* captured by the calibration. The analyst should evaluate these additional uncertainty sources independently, especially when the CN* sensitivity range is narrow relative to the HEC-HMS numerical precision.

9 Conclusions

This paper presented `gcm_scs_gradex_calib` v1.0.0, an open-source R module that operationalizes the GRADEX equal-slopes criterion for calibrating the SCS/HEC-HMS Curve Number as an automated, auditable and reproducible pipeline. The main contributions are:

(1) *End-to-end integration with GRADEX-DUAL v1.0*: α_P and its $\pm 30\,\%$ uncertainty range [3] are propagated through three calibration scenarios (optimistic, base, pessimistic).

(2) *Continuous CN estimator*: natural cubic spline and Brent minimization reduce discretization error from $O(\pm\Delta\text{CN}/2)$ to $O(10^{-3})$ CN units, with three robustness guards (G1–G3).

(3) *Comprehensive statistical validation*: OLS Gumbel-linearity ($R^2 \geq 0.95$), Shapiro-Wilk, Durbin-Watson and LOO-CV stability in a single automated workflow.

(4) *Bootstrap uncertainty*: non-parametric bootstrap ($n = 2\,000$) provides a distributional characterization of CN* independent of parametric assumptions.

(5) *Sensitivity and elasticity analysis*: structured α_P perturbation sweep yields the CN* sensitivity range and local elasticity directly usable in HEC-HMS sensitivity analysis.

(6) *LOO influence diagnostics*: identifies individual HEC-HMS simulations whose removal substantially shifts CN*, enabling targeted quality control.

(7) *Mathematical boundary of the method*: the Durbin-Watson autocorrelation warning observed in the case study does not invalidate the equal-slopes calibration. It reflects the inherent nonlinearity of the CN-to-ΔQ mapping that the OLS linearization approximates, and delineates a well-defined boundary of the method's applicability: the OLS Gumbel-linearity criterion ($R^2 \geq 0.95$) is met, but residual independence is not guaranteed. Analysts should inspect the residual pattern in Fig. 1 and report the Durbin-Watson result as part of the calibration documentation.

(8) *Automated refinement workflow*: when $E > 5\,\%$, the module computes the CN range and grid step required to achieve the tolerance threshold and exports a ready-to-use HEC-HMS simulation work list.

Acknowledgements

The author thanks the anonymous reviewers for their constructive comments, which substantially improved the quality of this manuscript.

Declarations

Conflict of interest. The author declares no conflict of interest.

Funding. This research was self-funded by the author.

Data availability. The case-study HEC-HMS simulation data and precipitation GRADEX inputs are embedded within the module via the `gcm_example_hec_data()` function and are fully reproducible by running `gcm_scs_gradex_calib.R` with `use_example = TRUE` in Step 2 of the execution pipeline (Section 12).

Code availability. `gcm_scs_gradex_calib` v1.0.0 is freely available under a CC BY 4.0 license at: https://github.com/MauricioVictoriaN/GCM-SCS-GRADEX-CALIB.

Companion module. GRADEX-DUAL v1.0 [3] is published at doi:10.31224/6945 and its R source code is available at: https://github.com/MauricioVictoriaN/GRADEX-DUAL.

References

[1] Guillot, P. and Duband, D. (1967). La méthode du Gradex pour le calcul de la probabilité des crues à partir des pluies. In *La prévision des crues et la protection contre les inondations, Dixièmes Journées de l'Hydraulique*, Rapport I.7. Société Hydrotechnique de France, Paris. https://www.persee.fr/doc/jhydr_0000-0001_1969_act_10_1_3763.

[2] Naghettini, M. (1995). O gradiente de chuvas intensas como parâmetro de projeto de obras hidráulicas. In *XI Simpósio Brasileiro de Recursos Hídricos*, Recife, Brazil.

[3] Victoria Niño, M.J. (2025). GRADEX-DUAL v1.0: Dual-method precipitation GRADEX estimator with L-moment regional frequency analysis. *engrXiv preprint*. doi:10.31224/6945. Code: https://github.com/MauricioVictoriaN/GRADEX-DUAL.

[4] Hydrologic Engineering Center (2016). *HEC-HMS Hydrologic Modeling System, User's Manual*. US Army Corps of Engineers, Davis, CA. https://www.hec.usace.army.mil/software/hec-hms/documentation.aspx.

[5] Soil Conservation Service (SCS) (1972). *National Engineering Handbook, Section 4: Hydrology*. USDA, Washington, D.C. https://ia801205.us.archive.org/27/items/CAT71334647003/CAT71334647003.pdf.

[6] Hosking, J.R.M. and Wallis, J.R. (1997). *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press. doi:10.1017/CBO9780511529443.

[7] de Boor, C. (2001). *A Practical Guide to Splines* (Revised ed.). Springer, New York. https://link.springer.com/book/9780387953663.

[8] Brent, R.P. (1973). *Algorithms for Minimization Without Derivatives*. Prentice-Hall, Englewood Cliffs, NJ. https://maths-people.anu.edu.au/~brent/pub/pub011.html.

[9] Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall, New York. Reprinted with corrections, Chapman & Hall/CRC, 1994. doi:10.1201/9780429246593.

[10] Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou, E. (1993). Frequency analysis of extreme events. In Maidment, D.R. (ed.), *Handbook of Hydrology*, Ch. 18. McGraw-Hill, New York. https://www.mheducation.com/highered/mhp/product/handbook-hydrology.html.

[11] Durbin, J. and Watson, G.S. (1950). Testing for serial correlation in least squares regression I. *Biometrika*, **37**(3–4), 409–428. doi:10.1093/biomet/37.3-4.409.

[12] Durbin, J. and Watson, G.S. (1951). Testing for serial correlation in least squares regression II. *Biometrika*, **38**(1–2), 159–177. doi:10.1093/biomet/38.1-2.159.

[13] Shapiro, S.S. and Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, **52**(3–4), 591–611. doi:10.1093/biomet/52.3-4.591.

[14] Zeileis, A. and Hothorn, T. (2002). Diagnostic checking in regression relationships. *R News*, **2**(3), 7–10. https://CRAN.R-project.org/doc/Rnews/Rnews_2002-3.pdf.

[15] Cook, R.D. (1977). Detection of influential observation in linear regression. *Technometrics*, **19**(1), 15–18. doi:10.2307/1268249.

[16] Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A. and Savenije, H. (eds.) (2013). *Runoff Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales*. Cambridge University Press. doi:10.1017/CBO9781139235761.

[17] Rigby, R.A. and Stasinopoulos, D.M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C*, **54**(3), 507–554. doi:10.1111/j.1467-9876.2005.00510.x.

[18] Wickham, H. and Bryan, J. (2023). *readxl: Read Excel Files*. R package version 1.4.3. https://CRAN.R-project.org/package=readxl.

[19] Ooms, J. (2023). *writexl: Export Data Frames to Excel "xlsx" Format*. R package version 1.4.2. https://CRAN.R-project.org/package=writexl.

[20] Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer, New York. doi:10.1007/978-3-319-24277-4.

[21] World Meteorological Organization (2008). *Guide to Hydrological Practices*, Vol. II (6th ed.). WMO-No. 168, Geneva. https://www.hydrology.nl/images/docs/hwrp/WMO_Guide_168_Vol_II_en.pdf.