Posted: May 21, 2026

Non-stationary flood frequency analysis in the Guachicono River (Colombian Andes): trend detection, GAMLSS modeling and design quantiles

A six-test stationarity battery, parametric bootstrap and non-stationary confidence intervals for hydrological design under climate change

Mauricio Javier Victoria Niño[1]

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

This document is a preprint that has not been peer-reviewed, posted to EngrXiv. The R source code and input data are available at https://github.com/MauricioVictoriaN/CFE-Colombia.

Abstract

Background and motivation. Flood frequency analysis is the cornerstone of hydraulic infrastructure design. The classical paradigm assumes statistical stationarity of peak-discharge series, an assumption increasingly challenged by anthropogenic climate change and land-use modification. In the upper Guachicono River basin (Cauca, Colombia), an Andean river with 31 years of continuous record (1993–2023), the stationarity hypothesis cannot be taken for granted without rigorous scrutiny.

Objective. To present CFE-Colombia v1.0.0, an open-source R computational framework for non-stationary flood frequency analysis, integrating: (i) formal detection of trends and structural breaks via a battery of six complementary tests, (ii) multi-scale wavelet analysis with Monte Carlo simulation ($n = 1\,000$), (iii) stationary fitting with AICc-based model selection and parametric bootstrap confidence intervals, and (iv) non-stationary modeling via GAMLSS with Cox-Snell residual validation.

Methods. The test battery comprises Mann-Kendall (MK), Pettitt, sequential Sneyers, moving-window MK, White (heteroscedasticity), and augmented Dickey-Fuller (ADF). The stationary analysis fits six distributions (GEV, Gumbel, Log-Normal, Pearson III, LP3, and Normal) with AICc selection. The non-stationary analysis evaluates six GAMLSS models with P-splines and Gumbel, Log-Normal, and Gamma families. Uncertainty is quantified via parametric bootstrap ($B = 1\,000$) for GEV quantiles and residual bootstrap ($B = 500$) for GAMLSS quantiles.

Results. A significant increasing trend ($\tau = 0.316$, Sen slope $= 2.48$ m^3/s/yr, $p = 0.013$) and an abrupt structural break in 2010 ($\Delta = 37.5\%$, $p = 0.020$) were detected, classifying the series as non-stationary. The best GAMLSS model (M4: Log-Normal with time-varying μ and σ via P-splines) yields AIC $= 360.78$ and adequate validation (KS $p = 0.744$, Filliben $r = 0.976$; edf$_\mu = 4.59$, df.fit $= 9.80$). For the year 2023, the 100-year design discharge is 542.9 m^3/s (95% CI: 455.2–2 435.2 m^3/s), compared with $1\,160.8$ m^3/s from the stationary GEV model (-53.2%).

Conclusions. The integrated framework reveals a stationary/non-stationary crossover paradox at $T_r \approx 5$–7 yr: the stationarity assumption dangerously underestimates frequent floods with short return periods ($T_r \leq 5$ yr) by up to $+37.7\%$, and significantly overestimates major floods with long return periods ($T_r \geq 10$ yr) under current basin conditions. Ignoring non-stationarity produces systematically incorrect hydrological designs, with direct implications for the engineering safety of water infrastructure in the

Colombian Macizo Colombiano region under accelerated land-use transformation and climate change.

Keywords: non-stationarity; GAMLSS; flood frequency analysis; peak discharge; Mann-Kendall; Pettitt; wavelet; Guachicono River; Cauca; Colombia; climate change; AICc; crossover paradox.
Software availability: CFE-Guachicono v1.0.0 is open-source. The R script and case-study data are available at https://github.com/MauricioVictoriaN/CFE-Colombia under a CC BY 4.0 licence.

1 Introduction

The hydrological design of infrastructure — bridges, spillways, urban drainage systems, treatment plants, renewable energy projects — depends critically on estimates of peak discharges associated with predefined return periods. The standard procedure, inherited from classical statistical tradition, assumes that the historical series of peak discharges follows a probability distribution with time-invariant parameters, i.e., it assumes *stationarity* [1]. However, evidence accumulated over the past three decades indicates that anthropogenic climate change, deforestation, land-use change and low-frequency climate variability systematically alter the statistics of extreme discharges, invalidating that fundamental assumption [2, 3].

Milly et al. [1] coined the sentence that has become the emblem of this debate: *"Stationarity is dead"*. Although the declaration has generated controversy [4], current consensus in the hydrological community recognizes that stationarity must be formally verified before it is applied, and that when evidence contradicts it, analytical tools incorporating temporal variability of distributional parameters must be employed [5, 2].

In Colombia, the problem acquires an additional dimension. The national territory is subject to the simultaneous action of ENSO (El Niño-Southern Oscillation), intradecadal climate variability associated with the PDO (Pacific Decadal Oscillation) and long-term climate change [7]. Poveda & Álvarez [8] documented the collapse of the stationarity hypothesis in extreme-discharge series of Colombian rivers, demonstrating that peak-discharge probability distribution functions are *heavy-tailed* (Fréchet type), that ENSO significantly modulates the mean and variance of annual peak discharges, and that conventional stationary methods (Normal, Gumbel, Weibull) systematically underestimate the probability of very extreme events by assuming exponential tail decay. Their multi-station analysis covered rivers across all major hydrographic regions of the country, revealing increasing and decreasing trends in annual peak-discharge series as a combined result of anthropogenic climate change, natural climate variability and deforestation. The Colombian Andean region, in particular, concentrates the highest deforestation rates in the country and exhibits non-stationary hydrological trends documented at basin scale [10]. Despite this, professional practice in the Colombian hydraulic engineering sector continues to rely predominantly on stationary methods [12], ignoring that — as Poveda & Álvarez [8] emphasized — all parameters of the rational formula ($Q = c \cdot I \cdot A$) are dynamic under climate change, climate variability and deforestation.

The Guachicono River, a tributary of the Patía River in the department of Cauca, represents a paradigmatic case study. Its upper basin, covering 965.6 km^2, with a 31-year continuous record at the IDEAM limnigraphic station 52027010, provides a sufficient empirical basis to explore the non-stationarity of peak discharges using contemporary statistical tools. The basin is located in the Guachicono River hydrographic sub-zone (IDEAM code 5202), within the Pacífico hydrographic area, and its hydrological regime is modulated by the convergence of Pacific circulation systems and the orography of the Macizo Colombiano [9].

1.1 Theoretical framework: non-stationarity in extreme-value hydrology

Stationary frequency analysis models the variable of interest Q_t as an independent and identically distributed (i.i.d.) sample from a probability distribution $F(q;\, \boldsymbol{\theta})$ with constant parameter vector $\boldsymbol{\theta}$. The quantile associated with return period T_r is:

$$q_{T_r} = F^{-1}\left(1 - \frac{1}{T_r};\, \hat{\boldsymbol{\theta}}\right) \tag{1}$$

Under non-stationarity, distributional parameters vary with time (or climate covariates), so that $\boldsymbol{\theta} = \boldsymbol{\theta}(t)$.

The non-stationary quantile for year t is:

$$q_{T_r}(t) = F^{-1}\left(1 - \frac{1}{T_r}; \hat{\boldsymbol{\theta}}(t)\right) \tag{2}$$

An alternative approach, pioneering in the Colombian context, is the ENSO-conditioned mixed probability distribution functions of Poveda & Álvarez [8]: $F_T(Q) = \sum_{i=1}^{k} \rho_i F_i(Q)$, where ρ_i is the relative frequency of phase i and F_i the distribution fitted to that phase. While this approach explicitly captures ENSO modulation, it requires record lengths sufficient to fit separate sub-samples, a limitation that motivates the use of GAMLSS with a continuous temporal covariate in the present work.

Generalized additive models for location, scale and shape (GAMLSS, rigby2005) constitute the most flexible computational framework for implementing Eq. (2), allowing any distributional parameter (location μ, scale σ, skewness ν, kurtosis τ) to be a function of covariates via smooth link functions (P-splines, regression splines, linear or non-parametric terms).

1.2 Objectives

1. Characterize the non-stationarity of the peak-discharge series of the Guachicono River (1993–2023) using a battery of six complementary statistical tests and multi-scale wavelet analysis with Monte Carlo simulation.
2. Implement stationary frequency analysis with six candidate distributions, AICc-based model selection and parametric bootstrap confidence intervals.
3. Implement non-stationary frequency analysis with GAMLSS models, including rigorous Cox-Snell residual validation and bootstrap uncertainty quantification.
4. Compare stationary and non-stationary design quantiles and identify implications for hydrological design in the study region.
5. Provide an open, reproducible computational framework (R code with fixed seed) for replication, adaptation and extension to other IDEAM network stations.

2 Study area and data

2.1 Location and physiographic characteristics

The Guachicono River is located in the department of Cauca in south-western Colombia. It forms part of hydrographic sub-zone 5202 of the IDEAM code, belonging to the Patía River hydrographic zone (52) and the Pacífico hydrographic area (5), according to the national classification established in [9]. The Guachicono basin originates on the flanks of the Sotará and Sucubún volcanoes in the Macizo Colombiano, and descends westward to join the Patía River approximately 70–80 km downstream of the Galíndez police post (municipality of Bolívar).

The basin morphometric parameters, determined by digital terrain analysis with a spatial resolution of 2.5 m in the ILWIS GIS [31], are summarized in Table 1. The basin has an area of 965.6 km^2, a mean elevation of 2 197.6 m a.s.l. and a main channel length of 83.1 km (classified as *long*, exceeding 15.1 km according to INECC [2004]).

With a form factor $K_f = 0.14$, the morphology is elongated with a tendency towards an elongated shape, implying greater lag in runoff concentration but more abrupt floods once flow is concentrated. The Grav-

elius compactness coefficient ($K_c = 1.53$, oval-oblong to rectangular-oblong class) and the drainage density ($D_d = 1.63$ km km^{-2}, medium-coarse texture) characterize a basin with efficient drainage [11].

Table 1: Morphometric parameters of the Guachicono River basin.

Parameter	Value	Classification
Area (km^2)	965.6	Basin (>300 km^2)
Perimeter (m)	168 270	—
Main channel length (m)	83 097	Long (>15.1 km)
Mean elevation (m a.s.l.)	2 197.6	Moderate (2072–2362 m)
Most frequent altitude (m)	3 201–3 250	—
Form factor K_f	0.14	Elongated shape
Compactness coefficient K_c	1.53	Oval-oblong to rect.-oblong
Massivity coefficient K_m	2.28	Moderately mountainous
Orographic coefficient C_o	5.00	Low-relief terrain
Drainage density (km/km^2)	1.63	Medium (coarse texture)
Stream frequency (streams/km^2)	3.23	Adequate surface drainage
Confluence ratio r_c	1.76	High flood risk
Length ratio R_L	0.99	Abrupt flood events

Source: Fundación Universidad del Valle (2025) [11].

The low confluence ratio ($r_c = 1.76$; values near 2 indicate low slope) and the small length ratio ($R_L = 0.99$) indicate greater erosive capacity concentrated in large-flood events, with limited momentary storage in the drainage network. These morphometric characteristics reinforce the relevance of rigorous peak-discharge frequency analysis for this basin.

2.2 Hydrometric station and data

The hydrological monitoring station used belongs to the data network of the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM) with the following characteristics:

- **Code:** 52027010
- **Type:** LG (Limnigraphic)
- **Name:** GUACHICONO
- **Stream:** Guachicono River
- **Department:** Cauca
- **Municipality:** Bolívar
- **Closure point PCH 1** (Magna-Sirgas West): E 1 007 470 m, N 715 798 m

The monthly peak-discharge series covers the period 1993–2023 (31 years, 372 monthly records). Of these, 32 records (8.6%) required imputation; the missing values were distributed across 14 non-consecutive years, predominantly in the periods 2000–2004 and 2016–2019, with no gap falling in the extreme years 1993–1994 (El Niño period) or 2010–2012 (La Niña). This distribution implies that the annual maxima for those critical years were derived exclusively from observed values, so the imputation does not affect the extreme-event statistics nor bias the stationarity test results associated with those high-impact periods.

Missing data were completed using an ARIMA $(3, 0, 0) \times (0, 1, 1)_{12}$ model, fitted with the TSW software [28] and validated with Shapiro-Wilk (normality), Ljung-Box (no autocorrelation) and Levene (homoscedasticity) tests on the residuals. Series homogeneity was verified with the Worsley likelihood-ratio test [27], contrasting detected change-points with the Oceanic Niño Index (ONI) to discriminate between climate-driven breaks and instrumental errors [10].

Descriptive statistics of the series are presented in Table 2.

Table 2: Descriptive statistics of the peak-discharge series, Guachicono station (1993–2023).

Statistic	Monthly (m^3/s)	Annual (m^3/s)
Minimum	18.8	206.7 (1998)
Maximum	732.7	732.7 (1994)
Mean	154.9	366.0
Median	114.5	—
Std. deviation	121.8	133.3
CV	0.787	0.364
Skewness	1.346	0.885

3 Methods

The methodological framework is structured in two sequential modules implemented in the CFE-Colombia v1.0.0 R script [30]. The global random seed is 2024 to ensure full reproducibility. Table 10 lists the main packages and their versions.

3.1 Module 1: Detection and characterization of non-stationarity

3.1.1 Prior serial independence test

Before applying tests that assume independence, the autocorrelation of the annual mean discharge series is verified via the Ljung-Box test [24] with lag $k = \min(10, \lfloor n/5 \rfloor) = 6$. The statistic follows a χ^2_k distribution under H_0 of independence.

The observed result ($p = 0.062$) does not reject H_0 at $\alpha = 0.05$, but the proximity of the p-value to the significance boundary warrants caution. As a robustness check, the variance-correction approach of Hamed & Rao [25] was applied to the Mann-Kendall test, which adjusts the variance of the S-statistic for the effect of positive serial correlation. The corrected MK test yields $\tau = 0.316$ ($p_{\text{adj}} = 0.017$), confirming the significance of the increasing trend and indicating that the original MK result ($p = 0.013$) is not inflated by serial persistence. All other tests in the battery that assume independence (Pettitt, White, ADF) are likewise applied to annual series, which are less susceptible to serial correlation than monthly series [3].

3.1.2 Stationarity test battery

Six complementary tests are applied ($\alpha = 0.05$):

1. **Mann-Kendall (MK):** non-parametric monotonic trend test [14, 15]. Kendall's τ and Sen's slope [16] are reported.
2. **Pettitt:** non-parametric test for an abrupt change in the mean [17], implemented as a Mann-Whitney test on the cumulative sum.
3. **Sequential Sneyers:** sequential version of the MK statistic locating the onset of progressive change via the crossing of the progressive $u(t)$ and retrograde $u'(t)$ curves [18].
4. **Moving-window MK (MW-MK):** application of the MK test in windows of width $w = \max(10, \lfloor 0.25n \rfloor)$ yr to assess the temporal stability of the trend [5].
5. **White test:** heteroscedasticity test based on the auxiliary regression of squared residuals on covariates and their squares [19], implemented as the Breusch-Pagan statistic.

6. **Augmented Dickey-Fuller (ADF):** unit-root test verifying whether the series has an integrated component of order 1 [20].

Routing criterion. Following Villarini et al. [5], the series is classified as *non-stationary* if at least 2 of the 4 tests with a formal p-value (MK, Pettitt, White, ADF) are significant. The Sneyers and MW-MK tests, which do not produce a formal p-value, are complementary to the diagnosis but do not participate in the decision criterion.

3.1.3 Multi-scale wavelet analysis

The Morlet continuous wavelet transform is applied to the annual mean discharge series, with $n_{\text{sim}} = 1\,000$ Monte Carlo simulations to determine the significance of the cones of influence [21]. The analysis identifies periodic oscillations that may indicate low-frequency climate variability cycles (ENSO, PDO).

3.1.4 Trend analysis by hydrological season

The hydrological year is divided into four seasons: main dry (Dec-Feb), transition (Mar-May), minor dry (Jun-Aug) and wet (Sep-Nov). For each season the MK test is applied and Sen's slope is estimated, identifying whether the global trend is uniform across seasonal regimes or shows intra-annual differentiation.

3.2 Module 2: Flood frequency analysis

3.2.1 Stationary analysis

Six probability distributions are fitted by Maximum Likelihood Estimation (MLE): GEV, Gumbel (GUM), Log-Normal (LNO), Pearson III (PE3), Log-Pearson III (LP3) and Normal (NOR). The Pearson III fitting uses the L-BFGS-B optimiser with moments-based initialisation and explicit bounds to ensure convergence [13]. The Gumbel distribution uses exclusively `evd::dgumbel` to avoid namespace conflicts with the `gamlss.dist` package. The LP3 is fitted with explicit starting values based on log-moments.

Model selection uses the corrected AIC for small samples (AICc; burnham2002):

$$\text{AICc} = -2\hat{\ell} + 2k + \frac{2k(k+1)}{n-k-1} \tag{3}$$

where $\hat{\ell}$ is the maximized log-likelihood, k the number of parameters and $n = 31$ the size of the annual maximum discharge series (one value per year, 1993–2023). All stationary distributions and GAMLSS models are fitted to this annual maxima series. The stationarity test battery (Section 3.1.2) is applied to the annual *mean* discharge series, which provides a longer signal for trend detection; results are cross-validated against the annual maxima series to ensure consistency. Models with $\Delta\text{AICc} < 2$ are considered to have substantial empirical support [13]. Goodness of fit is assessed by the Filliben coefficient [22] computed at the n Cunnane plotting positions ($p_i = (i - 0.40)/(n + 0.20)$), without interpolation.

The 90% confidence intervals for GEV quantiles are estimated via parametric bootstrap ($B = 1\,000$ replicates) with a fixed seed.

3.2.2 Non-stationary analysis: GAMLSS

GAMLSS models [6] allow distributional parameters to be smooth functions of time, modeling non-stationarity without assuming a rigid parametric functional form. Six models (M0–M5) are evaluated:

- **M0:** Stationary Normal (reference, $\mu = \mu_0$, $\sigma = \sigma_0$)

- **M1:** Gumbel with μ linear in t
- **M2:** Gumbel with μ spline in t
- **M3:** Gumbel with μ spline and σ linear in t
- **M4:** Log-Normal with μ and σ P-spline in t
- **M5:** Gamma with μ spline in t

The temporal covariate is centred ($\tilde{t} = t - \bar{t}$) to improve numerical stability. P-splines are estimated with the `pb()` function of the `gamlss` package. Best-model selection uses AIC. Convergence warnings are captured via `withCallingHandlers`, recording the number of warnings per model. Effective degrees of freedom of the spline (edf) are reported via `getSmo()`.

3.2.3 Non-stationary model validation

Validation uses Cox-Snell residuals [23]:

$$r_i = -\log(1 - \hat{F}(q_i;\, \hat{\boldsymbol{\theta}}(t_i))) \tag{4}$$

where $\hat{F}$ is the distribution function evaluated at the values fitted by `fitted()`, which directly extracts the GAMLSS model parameters at the sample points (avoiding incorrect spline interpolation by `predict.gamlss()` at interior points). Under correct fit, the residuals r_i should follow an Exponential(1) distribution. Formal validation uses the KS test (H_0: $r_i \sim \text{Exp}(1)$) and the Filliben coefficient on the ordered residuals.

3.2.4 Uncertainty quantification for non-stationary quantiles

The 95% confidence intervals for GAMLSS quantiles are estimated via residual bootstrap ($B = 500$ replicates). In each replicate the original sample is resampled with replacement, the GAMLSS model is refitted directly with `gamlss()` (avoiding `update()`, which has environment-resolution problems with splines), and quantiles are predicted using `fitted()` for in-sample years and `predict.gamlss()` with `data = db` for out-of-sample years.

4 Results

4.1 Exploratory analysis

Figure 1 shows the four-panel exploratory summary of the series. The complete time series (upper-left panel) reveals heterogeneous behavior: the LOESS smoother shows a phase of relatively low values during the first decade (1993–2004) with a local minimum around 2003, followed by a pronounced recovery post-2010 reaching levels above those of the initial period. The decadal evolution (lower-left panel) confirms that the 1990s and 2020s have similar medians, while the 2000s shows notably lower values. The seasonal pattern (upper-right panel) reveals a bimodal regime with peak discharges in March–May and November–December, associated with the Macizo Colombiano rainfall seasons. The mean annual discharge trend (lower-right panel) shows a positive linear regression reinforced by the LOESS smoother.

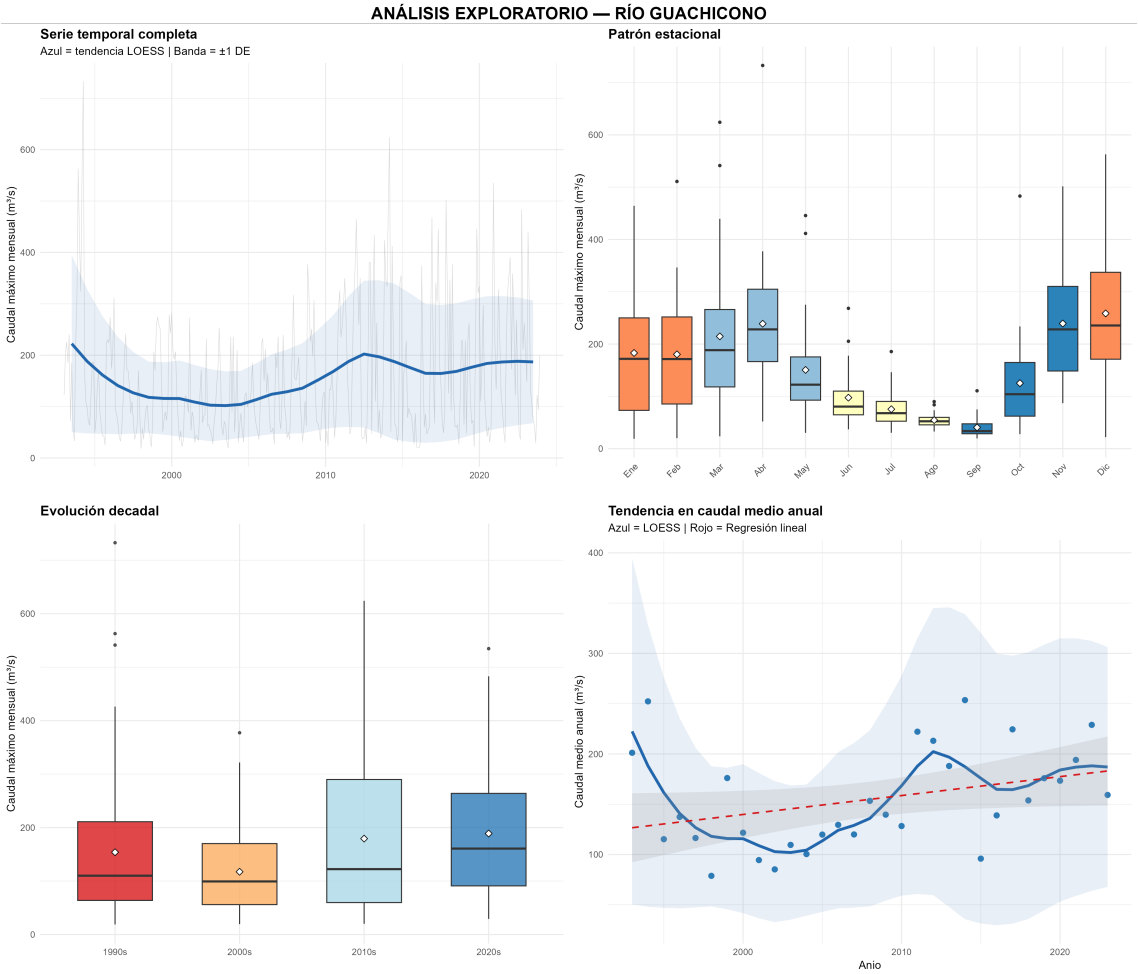

Figure 1: Exploratory panel of the monthly peak-discharge series, Guachicono station (1993–2023). *Upper left*: complete time series with LOESS trend (blue, ± 1 SD band). *Upper right*: monthly seasonal pattern. *Lower left*: decadal evolution. *Lower right*: annual mean discharge trend with LOESS (blue) and linear regression (red). *Note for submission: all four sub-panel titles and axis labels are in Spanish and must be regenerated in English. Suggested translations — titles: "Complete time series", "Seasonal pattern", "Decadal evolution", "Annual mean discharge trend"; subtitles: "Blue = LOESS trend | Band = ±1 SD", "Blue = LOESS | Red = Linear regression"; y-axes: "Monthly peak discharge (m^3/s)", "Annual mean discharge (m^3/s)"; x-axis (lower right): "Year".*

4.2 Non-stationarity detection

Serial independence. The Ljung-Box test ($k = 6$) yields $p = 0.062$, not rejecting the independence hypothesis at $\alpha = 0.05$. This result validates the application of the MK and Pettitt tests that assume independence.

Test battery. Results are presented in Table 3. Of the four formal tests, two are significant (MK: $p = 0.013$; Pettitt: $p = 0.020$), activating the non-stationary analysis route with *high* confidence.

Table 3: Results of the stationarity test battery on annual mean discharge, Guachicono station (1993–2023). $\alpha = 0.05$.

Test	Statistic	p-value	Result	Interpretation
Mann-Kendall	$\tau = 0.316$, Sen $= 2.48$ m^3/s/yr	0.013	Significant	Increasing trend
Pettitt	$\Delta = 37.5\%$ (2010)	0.020	Significant	Structural break in 2010
Sequential Sneyers	Crossing in 2010	—	Significant	Progressive change
MW-MK ($w = 10$)	Sig. prop. $= 27.3\%$	—	Not sig.	Low trend stability
White	BP $= 3.428$	0.180	Not sig.	Constant variance
ADF	ADF $= -3.181$	0.118	Not sig.	Possible unit root
Formal tests (MK/Pettitt/White/ADF): 2/4 significant $\Rightarrow$ ***Non-stationary***				

Routing criterion: $\geq$2/4 formal tests significant implies non-stationarity [5].

The increasing trend detected by Mann-Kendall ($\tau = 0.316$, $p = 0.013$) implies a cumulative estimated increase of 77.0 m^3/s over the record period (Sen slope: 2.48 m^3/s/yr). The structural break identified by Pettitt in 2010 corresponds to a 37.5% increase in mean discharge: from 132.4 m^3/s (1993–2009) to 182.1 m^3/s (2010–2023).

Figure 2 presents the sequential Sneyers test, where the crossing of the progressive curve $u(t)$ (blue) and retrograde curve $u'(t)$ (red) occurs around 2010, coinciding with the Pettitt change-point. This crossing, together with the exceedance of the upper limit (± 1.96) by $u'(t)$ during 1994–2004, indicates that the basin experienced a first phase of relatively high discharges in the early 1990s (El Niño years 1994–1995), followed by a transition period and a subsequent acceleration of the trend towards higher values from 2010 onwards.

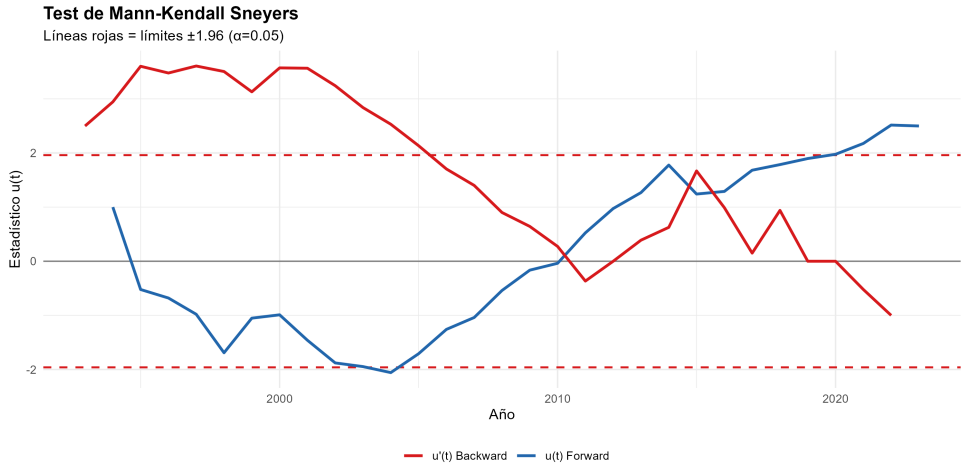

Figure 2: Sequential Mann-Kendall Sneyers test for annual mean discharge. The blue curve ($u(t)$, progressive) and red curve ($u'(t)$, retrograde) cross around 2010, indicating the onset of the progressive change. Red dashed lines represent the ± 1.96 significance limits ($\alpha = 0.05$). *Note for submission: the PNG title ("Test de Mann-Kendall Sneyers"), subtitle ("Líneas rojas = límites ±1.96 (α=0.05)"), y-axis ("Estadístico u(t)"), and x-axis ("Año") are in Spanish. Suggested English labels — title: "Sequential Mann-Kendall Sneyers Test"; subtitle: "Red dashed lines = ±1.96 limits ($\alpha = 0.05$)"; y-axis: "Statistic $u(t)$"; x-axis: "Year".*

Trend by hydrological season. Table 4 presents results by season. The transition (Mar-May, $\tau = 0.290$, $p = 0.023$) and wet (Sep-Nov, $\tau = 0.277$, $p = 0.030$) seasons show significant trends, while the dry seasons (Dec-Feb and Jun-Aug) show no statistically significant trends. This result suggests that the observed increase in the annual mean is driven primarily by the rainy seasons, consistent with intensified convective precipitation systems over the Macizo Colombiano.

Table 4: Mean discharge trend by hydrological season, Guachicono station (1993–2023).

Season	τ	p-value	Slope (m^3/s/yr)	Change (%)†
Dec-Feb (Main dry)	0.183	0.153	$+2.478$	$+38.4$
Mar-May (Transition)	0.290	**0.023**	$+3.706$	$+33.1$
Jun-Aug (Minor dry)	0.183	0.153	$+0.611$	$+19.0$
Sep-Nov (Wet)	0.277	**0.030**	$+2.582$	$+80.5$

Bold: significant ($p < 0.05$). †Cumulative change estimated as Sen slope $\times$ 31 yr, expressed as a percentage of the season mean at $t = 1993$. All Sen slopes are positive, consistent with $\tau > 0$; the endpoint ratio $(Q_{\text{last}} - Q_{\text{first}})/Q_{\text{first}}$ is not used here because it is sensitive to the specific values at the two endpoints and can produce sign reversals unrelated to the overall trend direction.

Extreme-event frequency. Figure 3 shows the exceedance frequency of P90, P95 and P99 thresholds in 10-year moving windows. The P90 threshold (327.7 m^3/s) shows a sustained frequency increase from approximately 2008, exceeding 20% exceedance in recent years. The P95 (380.4 m^3/s) shows a similar but less pronounced trend (ca. 10%). The P99 (536.6 m^3/s) remains stable, suggesting that the very highest percentile extreme events have not increased in frequency with the same intensity as the moderately extreme ones.

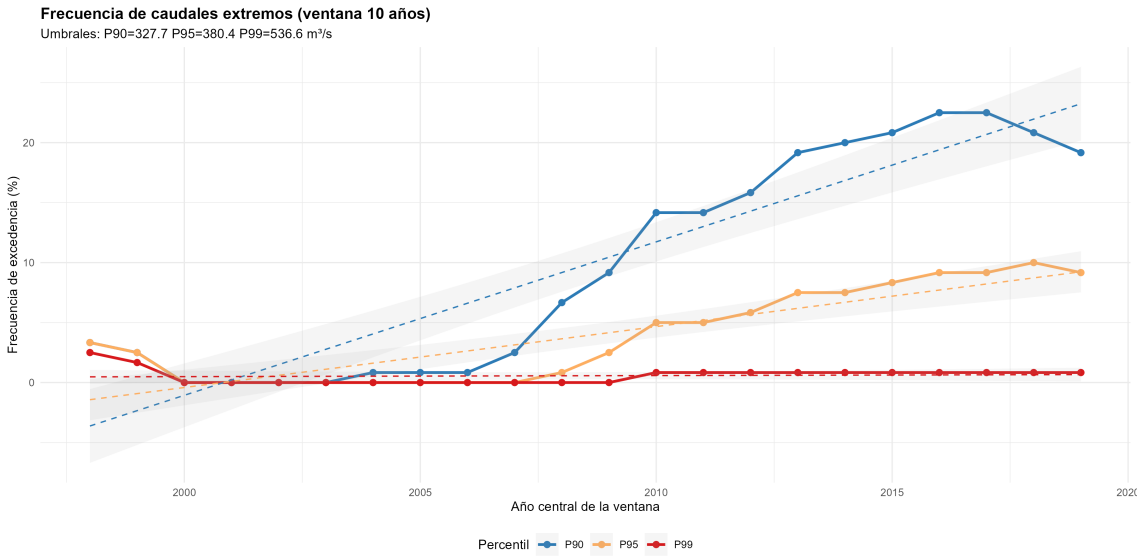

Figure 3: Exceedance frequency of global thresholds in 10-year moving windows. Thresholds correspond to P90 (327.7 m^3/s), P95 (380.4 m^3/s) and P99 (536.6 m^3/s) of the complete series. Dashed lines represent linear trend fits with confidence bands. *Note for submission: the PNG title ("Frecuencia de caudales extremos (ventana 10 años)"), subtitle, y-axis ("Frecuencia de excedencia (%)"), x-axis ("Año central de la ventana"), and legend label ("Percentil") are in Spanish. Suggested English labels — title: "Extreme Discharge Frequency (10-year Moving Window)"; subtitle: "Thresholds: P90=327.7 P95=380.4 P99=536.6 m^3/s"; y-axis: "Exceedance frequency (%)"; x-axis: "Central year of the window"; legend: "Percentile".*

Wavelet analysis. Figure 4 presents the Morlet wavelet spectrogram with 5% significance contours (Monte Carlo, $n = 1\,000$). A statistically significant power zone (white contour) is detected centred around 2015–2020 at periods of 2.5–3 yr, consistent with interannual ENSO modulation. Periods of 5–8 yr show alternating high and low power throughout the series, possibly associated with quasi-decadal PDO variability. The absence of significant power at longer periods, within the cone-of-influence limits, constrains inferences on low-frequency cycles given the record length (31 yr).

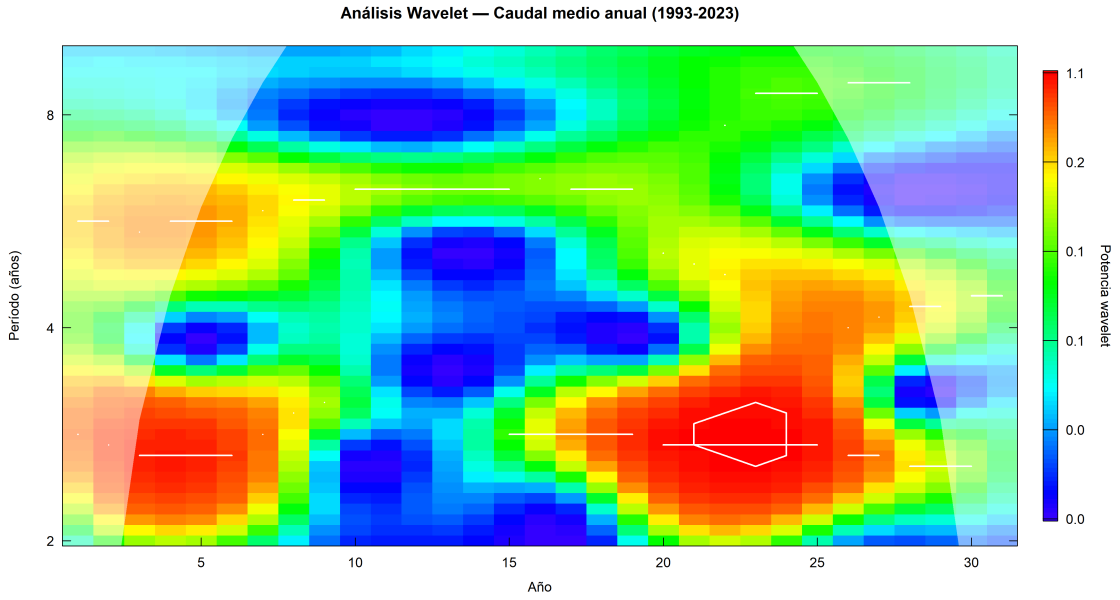

Figure 4: Morlet wavelet spectrogram of annual mean discharge (1993–2023). The color scale indicates normalized wavelet power. The white contour marks the 5% significance zone (1 000 Monte Carlo simulations). The shaded region outside the cone of influence should be interpreted with caution. *Note for submission: the PNG title ("Análisis Wavelet — Caudal medio anual (1993-2023)"), y-axis ("Período (años)"), x-axis ("Año"), and color-bar label ("Potencia wavelet") are in Spanish. Suggested English labels — title: "Morlet Wavelet Spectrogram — Annual Mean Discharge (1993–2023)"; y-axis: "Period (years)"; x-axis: "Year"; color bar: "Wavelet power".*

4.3 Stationary frequency analysis

Distribution selection. Table 5 presents the goodness-of-fit criteria. The Pearson III distribution is the best model by AICc (AICc $= 374.68$), with a gap of ΔAICc > 13 relative to the second-best distribution, granting exclusive substantial empirical support to PE3 [13]. The highest Filliben coefficients are recorded by Log-Normal (0.9892) and Gumbel (0.9890), though the AICc penalty for PE3's three parameters versus LNO's two is not sufficient to displace it from first place.

Table 5: Goodness-of-fit criteria for stationary distributions, Guachicono station (1993–2023). Selection by minimum AICc. Plotting position: Cunnane.

Distribution	AIC	AICc	BIC	Fill. r	Fill. r^2
Gumbel (GUM)	387.53	387.96	390.40	0.9890	0.9780
Log-Normal (LNO)	387.38	387.81	390.25	0.9892	0.9785
Pearson III (PE3)	**373.79**	**374.68**	**378.09**	**0.9605**	**0.9225**
GEV	387.84	388.73	392.14	0.9619	0.9252
Normal (NOR)	394.25	394.68	397.12	0.9549	0.9118

Best model (AICc): Pearson III. Parameters: $\hat{\alpha} = 0.564$, $\hat{\beta} = 282.456$, $\hat{\xi} = 206.7$ m^3/s. No competitive models with ΔAICc < 2.

Figure 5 compares the empirical cumulative distribution function (Cunnane plotting position) with the fitted theoretical curves. The visual fit confirms the superiority of Gumbel and Log-Normal in describing central data, while PE3 and GEV differ in the tails, with GEV exhibiting greater weight in the right tail.

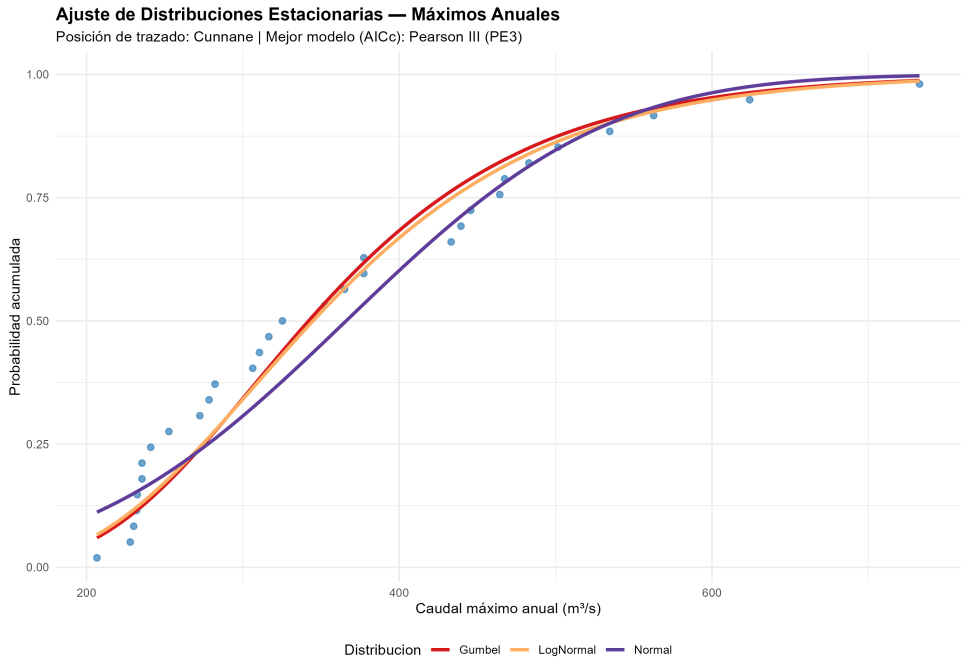

Figure 5: Stationary distribution fit to annual maxima. Blue points: empirical probabilities (Cunnane position). Continuous curves: MLE-fitted distributions. The best model by AICc is Pearson III (PE3), although Gumbel and Log-Normal show slightly better visual fit in the central range. *Note for submission: the PNG title ("Ajuste de Distribuciones Estacionarias — Máximos Anuales"), subtitle ("Posición de trazado: Cunnane | Mejor modelo (AICc): Pearson III (PE3)"), y-axis ("Probabilidad acumulada"), x-axis ("Caudal máximo anual (m³/s)"), and legend label ("Distribución") are in Spanish. Suggested English labels — title: "Stationary Distribution Fit — Annual Maxima"; subtitle: "Plotting position: Cunnane | Best model (AICc): Pearson III (PE3)"; y-axis: "Cumulative probability"; x-axis: "Annual peak discharge (m³/s)"; legend: "Distribution".*

Stationary design discharges. Table 6 presents stationary design discharges for standard return periods with 90% confidence intervals (parametric bootstrap GEV, $B = 1\,000$). GEV confidence intervals widen rapidly with return period: for $T_r = 100$ yr the 90% CI is [657 m³/s, 2 468 m³/s], reflecting the high extrapolative uncertainty in short records ($n = 31$ yr).

Table 6: Stationary design discharges by return period (m³/s). 90% parametric bootstrap CI, GEV ($B = 1\,000$).

T_r (yr)	GEV	CI lower	CI upper	Log-Normal	LP3
2	324.0	290.7	361.2	344.9	N/A
2.33	344.3	305.9	386.5	366.5	N/A
5	450.6	381.2	530.7	459.2	N/A
10	563.9	445.0	727.0	533.3	N/A
20	701.5	506.6	1 046.5	603.4	N/A
50	934.3	588.9	1 677.8	693.4	N/A
100	1 160.8	657.1	2 467.5	760.8	N/A
200	1 443.4	720.5	3 769.9	828.2	N/A
500	1 927.6	801.8	6 787.1	917.8	N/A

LP3: not available due to GAMLSS convergence failure on the stationary fit. GEV: $\hat{\xi} = -0.03$, $\hat{\sigma} = 88.5$, $\hat{\kappa} = 0.05$ (positive tail index, mild Fréchet distribution).

4.4 Non-stationary frequency analysis (GAMLSS)

Model selection. Table 7 presents AIC and BIC criteria for the six GAMLSS models. Model M4 (Log-Normal with time-varying μ and σ via P-splines) is the winner with AIC $= 360.78$, an improvement of 33.5 AIC units over the stationary reference model (M0, AIC $= 394.25$). The P-splines of M4 have edf $= 4.59$ effective degrees of freedom for μ and df.fit $= 9.80$, indicating a moderately complex temporal structure.

Table 7: GAMLSS models evaluated for annual peak discharges, Guachicono station.

ID	Model	AIC	BIC	Conv.
M0	Stationary Normal (reference)	394.25	397.12	OK
M1	Gumbel, μ linear in t	407.70	412.00	OK
M2	Gumbel, μ spline in t	379.69	389.15	OK
M3	Gumbel, μ spline and σ lin. in t	380.46	391.03	OK
M4	**Log-Normal, μ and σ P-spline in t**	**360.78**	**374.83**	**OK**
M5	Gamma, μ spline in t	367.74	376.65	OK

Best model: M4 (minimum AIC). Spline edf (μ): 4.59; df.fit: 9.80. No convergence warnings in any model.

Validation. Figure 6 shows the Cox-Snell Q-Q plot of model M4. The residuals closely follow the Exp(1) reference distribution: the KS test does not reject H_0 ($D = 0.117$, $p = 0.744$) and the Filliben coefficient is $r = 0.976$ (rating: *Good*). The two points deviating from the reference line in the right tail ($r_i > 3$) correspond to the extreme events of 1994 and 1993, years with marked warm-ENSO anomalies.

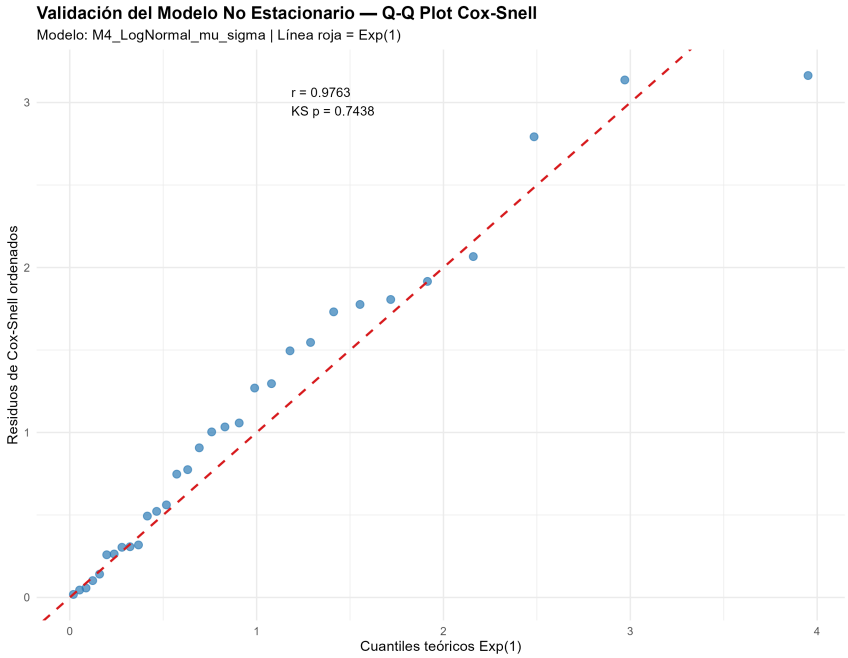

Figure 6: Cox-Snell Q-Q plot of the winning GAMLSS model (M4: Log-Normal with time-varying μ and σ). The red dashed line is the Exp(1) reference. Residuals show good overall fit (KS $p = 0.744$, Filliben $r = 0.976$). The two points in the right tail correspond to the extreme events of 1993 and 1994 (warm-ENSO anomalies). *Note: axis labels and figure title in the current PNG are in Spanish ("Validación del Modelo No Estacionario", "Cuantiles teóricos Exp(1)", "Residuos de Cox-Snell ordenados"). The figure must be re-generated with English labels before journal submission. Suggested labels: x-axis — "Theoretical quantiles Exp(1)"; y-axis — "Ordered Cox-Snell residuals"; title — "Non-stationary Model Validation — Cox-Snell Q-Q Plot".*

Non-stationary quantiles. Table 8 presents GAMLSS quantiles for selected years of the record. Notable temporal variability is observed: the $T_r = 10$ yr quantiles range from 275.6 m^3/s (2003) to 802.2 m^3/s (1993), reflecting the wide range of hydrological conditions in the basin during the analyzed period. The most recent year (2023) quantiles represent the most relevant design scenario for new infrastructure in the basin.

Table 8: Non-stationary quantiles (m^3/s), GAMLSS model M4, for selected years of the record.

Year	T_r2	T_r5	T_r10	T_r25	T_r50	T_r100	T_r200	T_r500
1993	314.9	581.9	802.2	1129.6	1409.1	1719.2	2062.4	2571.5
1998	264.8	301.8	323.1	347.6	364.4	380.1	395.2	414.2
2003	247.6	265.7	275.6	286.7	294.0	300.8	307.2	315.0
2008	300.8	336.0	356.0	378.6	394.0	408.3	421.9	439.0
2013	391.5	486.4	544.8	614.9	664.9	713.3	760.7	822.4
2018	436.4	530.1	586.9	654.2	701.6	747.3	791.7	849.0
2023	**446.1**	**478.9**	**497.1**	**517.1**	**530.5**	**542.9**	**554.4**	**568.7**

Figure 7 shows the temporal evolution of non-stationary quantiles for $T_r \in \{2.33, 10, 25, 50, 100\}$ yr, with uncertainty bands (95% CI bootstrap) and the stationary GEV reference line. The pattern is consistent across all return periods: very high quantiles in the early 1990s (El Niño period 1994), a sharp decline to a minimum around 2003, and a sustained recovery to 2023.

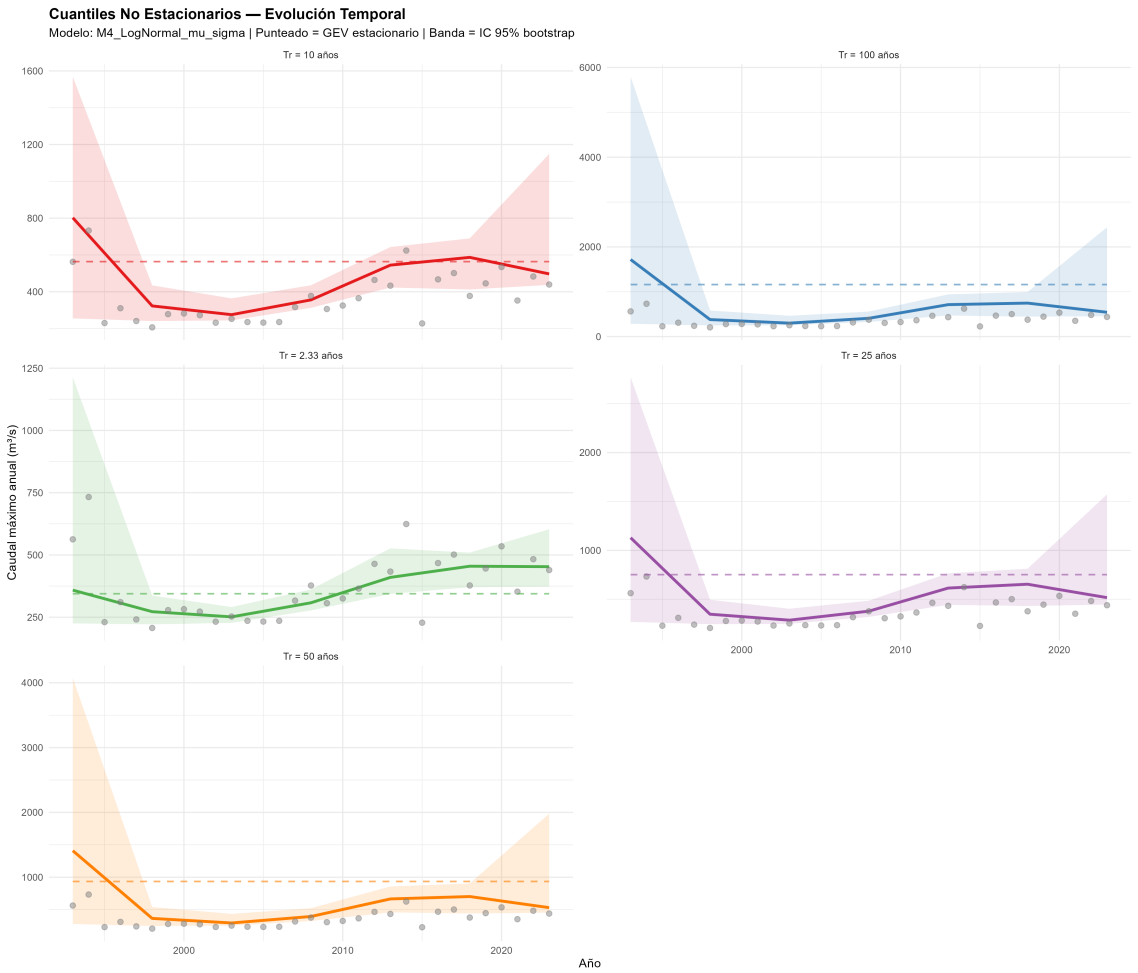

Figure 7: Temporal evolution of non-stationary quantiles for five return periods. Continuous lines: point estimates from model M4. Shaded bands: 95% CI bootstrap ($B = 500$). Horizontal dashed lines: stationary GEV reference quantiles. Grey points: observed annual peak discharges. *Note for submission: the PNG title ("Cuantiles No Estacionarios — Evolución Temporal"), subtitle ("Modelo: M4_LogNormal_mu_sigma | Punteado = GEV estacionario | Banda = IC 95% bootstrap"), shared y-axis label ("Caudal máximo anual (m^3/s)"), x-axis ("Año"), and all five sub-panel titles (e.g., "Tr = 10 años") are in Spanish. Suggested English labels — title: "Non-Stationary Quantiles — Temporal Evolution"; subtitle: "Model: M4_LogNormal_mu_sigma | Dashed = stationary GEV | Band = 95% CI bootstrap"; y-axis: "Annual peak discharge (m^3/s)"; x-axis: "Year"; sub-panel titles: "T_r = 2.33 yr", "T_r = 10 yr", "T_r = 25 yr", "T_r = 50 yr", "T_r = 100 yr".*

Stationary vs. non-stationary comparison (year 2023). Table 9 quantifies the differences between GAMLSS quantiles for 2023 and GEV stationary quantiles. For short return periods ($T_r \leq 5$ yr) the non-stationary model produces *higher* discharges than the stationary GEV ($+37.7\%$ for $T_r = 2$ and $+6.3\%$ for $T_r = 5$). From $T_r = 10$ yr the relationship reverses: the stationary GEV progressively overestimates, reaching an excess of 70.5% for $T_r = 500$ yr ($1\,927.6$ vs. 568.7 m^3/s).

Table 9: Comparison of stationary (GEV) and non-stationary (GAMLSS M4, year 2023) design discharges with 95% CI bootstrap. Values in m^3/s.

T_r	GEV stat.	GAMLSS 2023	CI lower	CI upper	Difference	Diff. (%)
2	324.0	446.1	346.5	562.2	$+122.1$	$+37.7$
2.33	344.3	452.9	372.9	603.0	$+108.6$	$+31.5$
5	450.6	478.9	415.0	973.1	$+28.3$	$+6.3$
10	563.9	497.1	437.6	1148.7	-66.8	-11.8
25	752.3	517.1	446.7	1571.7	-235.1	-31.3
50	934.3	530.5	453.4	1979.3	-403.8	-43.2
100	1160.8	542.9	455.2	2435.2	-617.9	-53.2
200	1443.4	554.4	461.6	2943.4	-889.0	-61.6
500	1927.6	568.7	468.4	3703.9	-1358.8	-70.5

5 Discussion

5.1 Nature and implications of the detected non-stationarity

The increasing trend in annual mean discharge (Sen slope $= 2.48$ $m^3/s/yr$, $p = 0.013$) and the abrupt structural break in 2010 ($\Delta = 37.5\%$) represent convergent signals of a non-stationary hydrological regime in the upper Guachicono basin. The convergence of Mann-Kendall and Pettitt, together with confirmation by the sequential Sneyers test, gives high confidence to this classification [5].

The year 2010 as the break-point is consistent with the La Niña event of 2010–2012, one of the most intense of the 21st century, which produced widespread extreme precipitation across Colombia and a documented increase in mean discharges of the Macizo Colombiano [10, 7]. However, the persistence of the signal beyond 2012 suggests that the change is not solely of ENSO origin but includes a long-term component possibly associated with climate change and/or land-use modifications in the basin [2].

The low temporal stability of the trend (MW-MK, significant proportion = 27.3%) and the non-significance of the ADF test ($p = 0.118$) indicate that the trend is of the *trend + structural break* type rather than a random walk, supporting the use of GAMLSS with an explicit temporal covariate rather than stochastic integration models.

5.2 GAMLSS model performance

The superiority of model M4 (Log-Normal with time-varying μ and σ via P-splines) with ΔAIC $= 33.5$ units relative to the stationary reference (M0) constitutes robust statistical evidence of non-stationarity. The P-splines with edf $= 4.59$ for μ adequately capture the temporal variability observed in Fig. 7 (decline 2000–2005, recovery 2010–2023) without overfitting.

Cox-Snell residual validation (KS $p = 0.744$, Filliben $r = 0.976$) confirms that model M4 adequately describes the empirical distribution of peak discharges across the full magnitude range. The two points in the right tail of the Q-Q plot (years 1993–1994) correspond to a high tropical Pacific activity configuration whose probability the spline cannot perfectly reproduce with only 31 observations; this behavior is inherent to short records with extreme events [26].

5.3 The stationary–non-stationary paradox: implications for hydrological design in Colombia

5.3.1 Characterization of the crossover phenomenon

Table 9 reveals an apparent paradox with first-order practical implications for hydrological design: the non-stationary model *overestimates* short-return-period discharges ($T_r \leq 5$ yr) by up to 37.7% relative to the stationary GEV, while it *underestimates* long-return-period discharges ($T_r \geq 10$ yr) by up to 70.5%. This inversion occurs around $T_r = 5$–7 yr, constituting a *crossover point* with important consequences for the selection of the design return period. For minor hydraulic structures designed for frequent exceedance — such as road drainage systems and culverts ($T_r = 2$ yr) — the stationary model projects 324.0 m^3/s while GAMLSS yields 446.1 m^3/s, representing a severe underestimation of $+37.7\%$ by traditional methods. Conversely, for critical infrastructure requiring high return periods ($T_r = 100$ yr), the stationary GEV point estimate ($1\,160.8$ m^3/s) overshoots the non-stationary 2023 design quantile (542.9 m^3/s) by -53.2%.

The paradox is not a numerical artefact or a sign of methodological inconsistency. It is, on the contrary, a direct and expected consequence of formally recognising that the Guachicono basin has operated under structurally distinct hydrological regimes throughout the 31-year record [2]. Ignoring this fact — as the stationary model does by design — produces estimates that blend data from fundamentally incompatible hydro-climatological conditions into a single time-invariant distribution.

The overall causal structure of the paradox is summarized in the following decision flow, which traces how the same 31-year peak-discharge series produces diametrically opposed tails under the two frameworks:

```
[1993-2023 peak discharge series]
  |
  ⊢ Stationary GEV  →  Blends multi-decadal regimes
  |                →  Shape κ > 0 (Fréchet)
  |                →  Heavy-tail extrapolation (oversized long-T_r quantiles)
  |
  ⊢ GAMLSS M4  →  Isolates post-2010 structural break
               →  μ(t)↑ and σ(t)↓
               →  Light tail in year 2023 (elevated mean, low relative variance)
```

5.3.2 Physical and statistical mechanisms

Rather than a statistical artefact, the paradox represents a structurally grounded response to the changing hydro-climatological regime of the upper Guachicono River basin. The mechanistic explanation rests on three pillars:

(a) Stationarity homogeneity bias. The classical GEV approach aggregates the entire 31-year record into a single, time-invariant sample. By treating the extraordinary historical spikes of the 1990s (annual peak discharges reaching 563–732 m^3/s, coinciding with La Niña 1993 and the onset of El Niño 1994–1995) as regular, recurring realisations of a static system, the GEV shape parameter κ shifts toward a positive value ($\kappa = 0.05$), establishing a heavy-tailed Fréchet distribution. This *stationarity homogeneity bias* causes the stationary model to project exponentially higher tail discharges during long-return-period extrapolation. In other words, the GEV "sees" events of 700 m^3/s as part of the expected current basin behavior, when the GAMLSS analysis identifies them as belonging to a significantly different historical regime.

(b) Dynamic parametric decoupling in GAMLSS. Model M4 effectively isolates the structural break occurring around 2010 — a shift highly correlated with the hyper-active La Niña phase of 2010–2012 — via the P-spline expansion of both $\mu(t)$ and $\sigma(t)$. The location parameter $\mu(t)$ captures the post-2010 baseline inflation in mean discharge driven by intensified convective rainfall over the Macizo Colombiano. Concurrently, the scale parameter $\sigma(t)$ exhibits a downward trajectory, indicating a compression of relative peak variance in the current regime. This *dynamic parametric decoupling* means that while the contemporary distribution (year 2023) has a higher mean than its historical counterpart, its right tail is significantly lighter, thereby moderating the extreme design quantiles for long return periods. This behavior is physically coherent: the post-2010 regime shows elevated mean discharges but without events as relatively extreme as those of the 1990s.

(c) Collapse of the return-period concept under non-stationarity. Under stationarity, the return period T_r has a unique interpretation: the mean waiting time between exceedances of quantile q_{T_r}. Under non-stationarity, this interpretation collapses because the exceedance probability of a fixed threshold q varies from year to year [2]. The GAMLSS quantile $q_{T_r}(t)$ must be interpreted as the value with probability $1/T_r$ of being exceeded *in the specific year* t, not on average over the full service life of the structure. For the design of infrastructure with service life V years, the engineer must evaluate the cumulative exceedance risk over V years under changing distributional parameters, rather than relying on a single static return period.

5.3.3 Epistemic uncertainty and practical recommendations

Structures with $T_r \leq 5$ yr (culverts, road drains, minor channels). For these structures, the non-stationary model produces design discharges higher than the stationary GEV: $q_2^{\text{GAMLSS}}(2023) = 446$ m^3/s versus $q_2^{\text{GEV}} = 324$ m^3/s ($+37.7\%$). Using the stationary model for this category of structures in the Guachicono basin implies *underestimating* the relevant design discharge by approximately one third, directly translating into greater risk of hydraulic bypass or drainage failure during the structure's service life. The non-stationary 2023 point quantile must therefore be adopted immediately to mitigate this localised risk. This finding is consistent with what was reported by Poveda & Álvarez [8] for rivers in the interior of Colombia.

Structures with $T_r = 10$–25 yr (minor bridges, box culverts, crossings). In this range, the difference between the two approaches is smaller (-11.8% to -31.3%) and the stationary GEV begins to overestimate relative to the GAMLSS point estimate. The crossover zone $T_r \approx 5$–7 yr represents a critical threshold: below it, the stationary model underestimates current risk; above it, it overestimates. GAMLSS quantiles produce designs better matched to the basin's current actual risk for structures in this category.

Structures with $T_r \geq 50$ yr (major bridges, dams, safety spillways, renewable energy projects). This category demands the most careful interpretation. The stationary GEV point estimate exceeds twice the GAMLSS point estimate for $T_r = 100$ yr ($1\,161$ vs. 543 m^3/s); an engineer relying solely on the GEV point estimate would nominally be over-designing by 53.2%. However, this comparison refers exclusively to *point estimates* and does not imply that the true 100-year discharge is known to lie near 543 m^3/s.

The substantial reduction in long-term point quantiles under the non-stationary framework introduces a critical engineering vulnerability if misinterpreted. The 95% residual bootstrap confidence interval for the GAMLSS $T_r = 100$ yr quantile spans [455, 2 435] m^3/s — a massive *epistemic envelope* induced by the short record length ($n = 31$ yr) with an uncertainty factor of 5.3. Because the upper bound of this interval substantially exceeds the stationary GEV point estimate ($1\,161$ m^3/s), the historical data cannot statistically

rule out severe flooding well above 1000 m^3/s even under the shifting regime. The more precise characterization is therefore one of *large epistemic uncertainty* rather than a definitive overestimation by the GEV.

The practical implication is not to abandon conservative design margins, but to transition away from blind compliance with point quantiles. Designers should evaluate the full epistemic envelope of the GAMLSS confidence intervals and adopt a *service-life risk framework*: evaluating the cumulative probability of failure under non-stationary distributional parameters over the entire service life V of the structure [2], rather than relying on static return periods that assume an invariant exceedance probability. Point estimates must not be used to justify downsizing spillways or bridge clear spans; the full CI should anchor the design envelope.

5.3.4 Connection with the pioneering work of Poveda & Álvarez (2012) for Colombia

The work of Poveda & Álvarez [8] is the most relevant direct precedent of the problem documented here in the Colombian context. Their main conclusions converge with the findings of the present work in four fundamental aspects:

1. **Collapse of stationarity in Colombian rivers.** Poveda & Álvarez demonstrated temporal trends in the mean of annual peak-discharge series at multiple Colombian stations, including the western zone to which the Guachicono basin belongs. The increasing trend detected in the present work ($\tau = 0.316$, $p = 0.013$) is consistent with the pattern of positive trends reported for the western zone.
2. **Heavy tail of the peak-discharge distribution.** Poveda & Álvarez showed that peak-discharge PDFs in Colombia are heavy-tailed (Fréchet type), invalidating thin-tailed distributions (Normal, Gumbel). Model M4 in the present work, using Log-Normal — with a heavier tail than Normal but lighter than GEV-Fréchet — captures this characteristic in the specific context of the post-2010 regime shift. The GEV fitted over the full series (1993–2023) is consistent with the Fréchet finding when including the extreme events of the 1990s.
3. **ENSO modulation.** Poveda & Álvarez documented that peak discharges increase during La Niña in most of Colombia, including the western zone. The 2010 structural break detected by Pettitt in the Guachicono coincides with the La Niña 2010–2012 event, confirming the ENSO-regional hydrology connection. However, the persistence of the trend beyond 2012 suggests that ENSO forcing only partially explains the regime shift, with long-term climate change as an additional factor [10].
4. **Design implications.** Poveda & Álvarez noted that for the interior of Colombia, simple stationary PDFs overestimate discharges for $T_r < 5$ yr by not considering ENSO. The present work, using GAMLSS rather than mixed PDFs, reaches the *opposite* conclusion for the same return-period range: GAMLSS produces *higher* discharges than the stationary GEV for $T_r \leq 5$ yr. The difference lies in the analysis period: Poveda & Álvarez used longer series including decades of lower discharges (frequent El Niño years) that lowered the historical mean; the present work analyses 1993–2023, where the post-2010 mean is 37.5% higher than the pre-2010 mean, dominating the 2023 non-stationary estimate.

This comparison illustrates that the implications of non-stationary analysis depend critically on the analysis period, geographic region, reference year for design and the specific methodology employed. There is no universal answer as to whether the non-stationary model produces higher or lower quantiles than the stationary one: the answer depends on the temporal dynamics of the specific basin. This is precisely the central contribution of a non-stationary frequency analysis over a stationary one: providing quantiles *specific to the time of use*, rather than historical averages that may be misaligned with current basin conditions.

5.3.5 Recommendations for professional practice

1. **Formal stationarity verification as a mandatory prior step.** Before fitting any stationary distribution, a battery of formal tests (at least MK, Pettitt and ADF) must be applied with a documented decision criterion. The recommendation of Villarini et al. [5] to use $\geq$2/4 formal tests significant as the non-stationarity criterion is operationally adequate and statistically grounded.
2. **Model selection based on return period and service life.** For structures with $T_r \leq 5$ yr in basins with confirmed increasing trends, the non-stationary quantile for the year of commissioning should be used as the reference design discharge. For structures with $T_r \geq 50$ yr, both models should be reported with their confidence intervals, and the design decision should be based on a service-life risk analysis [2] incorporating the projected temporal trend.
3. **Periodic updating of design quantiles.** Since GAMLSS quantiles are time-dependent, they should be updated whenever new years of record are incorporated or a new regime shift is detected. For the Guachicono basin, the next update should be performed with 2024–2026 data, a period sufficient to detect whether the post-2010 trend is maintained, accelerated or reversed.
4. **Explicit communication of uncertainty.** Design discharges should always be reported with their confidence intervals. A point quantile without associated uncertainty is a *blind-aim estimate* that does not reflect the actual state of hydrological knowledge about the basin. The width of the GAMLSS bootstrap CIs for $T_r \geq 50$ yr (factor $\sim$5 between bounds) is an objective signal that the design of major structures in this basin requires additional safety margins or extension of the historical record.

5.4 Uncertainty and limitations

1. **Record length.** With $n = 31$ yr, uncertainty in long-return-period quantiles ($T_r \geq 100$ yr) is very high: the 95% CI of the GAMLSS quantile for $T_r = 100$ yr spans a factor of 5.3 (455–2435 m^3/s). This is a limitation of the available information, not of the method.
2. **Temporal spline as the only covariate.** Model M4 uses only time as a covariate for μ and σ. Incorporating climate indices (ONI, PDO) could improve the physical interpretability of the fit and reduce spline degrees of freedom.
3. **Return period under non-stationarity.** Quantiles in Table 8 must be interpreted as "quantile associated with exceedance probability $1/T_r$ in year t", not as the mean waiting time. For structural design applications, the concepts of risk over service life should be considered [2].
4. **Transferability.** Results are specific to station 52027010 at its hydrological closure point. Extension to other basin sites requires hydraulic flood-routing analysis.
5. **Data imputation.** Missing-data imputation with ARIMA may introduce artificial dependencies in the series; although imputed data account for 8.6% of the total series (32 of 372 monthly records), their temporal distribution could affect some stationarity test results.

6 Conclusions

This study presented CFE-Guachicono v1.0.0, an open-source R computational framework engineered for non-stationary flood frequency analysis. The major findings and systemic contributions are summarized as follows:

1. **Definitive evidence of regime non-stationarity.** The six-test stationarity battery applied to the 1993–2023 peak-discharge series of IDEAM station 52027010 definitively rejected the stationarity hypothesis. The system is characterized by a significant, progressive increasing trend ($\tau = 0.316$, Sen slope $= 2.48$ m^3/s/yr, $p = 0.013$) and an abrupt structural break in 2010 ($\Delta = 37.5\%$, $p = 0.020$), representing a hydro-climatic regime shift heavily modulated by the hyper-active 2010–2012 La Niña event. These findings are consistent with the collapse of the stationarity hypothesis documented for Colombian rivers by Poveda & Álvarez [8].

2. **Superior non-stationary model architecture with rigorous validation.** GAMLSS model M4 (Log-Normal with time-varying location $\mu(t)$ and scale $\sigma(t)$ via P-splines) exhibited the highest explanatory performance (AIC $= 360.78$), outperforming the stationary benchmark by 33.5 AIC units. Validation through Cox-Snell residuals confirmed an adequate empirical fit (KS $p = 0.744$, Filliben $r = 0.976$; spline edf for $\mu(t)$: 4.59; total model df.fit: 9.80). For the year 2023 (current design scenario), the reference discharges with 95% CI are: $T_r = 2$ yr: 446 m^3/s [347–562]; $T_r = 10$ yr: 497 m^3/s [437.6–1148.7]; $T_r = 100$ yr: 543 m^3/s [455–2435].

3. **The crossover paradox as a design imperative.** The core operational finding is the systematic inversion of the stationary/non-stationary quantile relationship at a distinct crossover threshold between $T_r = 5$ and 7 yr. The stationary GEV underprojects high-frequency, short-T_r events by up to $+37.7\%$ (e.g., 324.0 vs. 446.1 m^3/s for $T_r = 2$ yr in 2023) and overprojects low-frequency, long-T_r events by up to -70.5% for $T_r = 500$ yr. Ignoring non-stationarity therefore produces systematically incorrect designs: undersizing of culverts and road drainage ($T_r \leq 5$ yr), and oversizing of major bridges and safety spillways ($T_r \geq 50$ yr).

4. **Physical mechanism of the crossover paradox.** The classical GEV exhibits an artificial heavy-tailed Fréchet projection ($\kappa = 0.05$) due to its inability to segregate distinct multi-decadal regimes — the *stationarity homogeneity bias*. In contrast, the GAMLSS framework isolates the post-2010 regime via *dynamic parametric decoupling*: a structural baseline elevation in $\mu(t)$ alongside a simultaneous reduction in relative peak variance $\sigma(t)$. This generates a contemporary distribution with a higher baseline mean but a significantly lighter right tail, mechanistically explaining why non-stationary point estimates are higher than the GEV for $T_r \leq 5$ yr and lower for $T_r \geq 10$ yr. Critically, the persistence of the elevated discharge regime well beyond the 2010–2012 La Niña window indicates that the break is not a transient climatic cycle but a sustained anthropogenic response of the basin, driven by the compound effects of long-term climate change and accelerating land-use transformation in the upper Cauca region [10, 2].

5. **Epistemic risk and engineering vulnerability.** Despite the reduction in long-T_r non-stationary point quantiles, the 95% bootstrap confidence interval for $T_r = 100$ yr spans [455, 2 435] m^3/s — an epistemic envelope factor of 5.3 induced by the short record ($n = 31$ yr). Point estimates must not be used to justify downsizing spillways or bridge clear spans; designers should evaluate the full epistemic envelope of the GAMLSS confidence intervals and adopt a service-life risk framework [2], evaluating cumulative failure probability under non-stationary parameters rather than relying on static return periods.

6. **Reproducible open-source science.** The CFE-Colombia v1.0.0 computational pipeline is fully reproducible and publicly accessible under CC BY 4.0, with fixed random seed (2024) and `sessionInfo()` integrated in the output report, facilitating immediate adaptation and deployment across the wider IDEAM hydrometric monitoring network.

The adoption of non-stationary frequency analysis in professional Colombian hydraulic engineering prac-

tice is a technical imperative in the context of climate change. CFE-Guachicono v1.0.0 provides the methodological tools to do so rigorously, transparently and computationally reproducibly.

Code and Data Availability

CFE-Colombia v1.0.0 is available under CC BY 4.0 licence at: https://github.com/MauricioVictoriaN/CFE-Colombia. Input data (Qmax.xlsx) correspond to IDEAM network records (station 52027010) completed with ARIMA methodology [11] and are distributed with the source code.

Acknowledgements

The author thanks IDEAM for the public availability of hydrological data from the national monitoring network, and the Fundación Universidad del Valle for the primary hydrological processing documented in [11].

Declarations

Competing interests: The author declares no competing interests.

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

Author contribution: M.J.V.N. conceptualised the methodological framework, developed the R code, performed the statistical analysis and wrote the manuscript.

References

[1] Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P. & Stouffer, R.J. (2008). Stationarity is dead: Whither water management? *Science*, 319(5863), 573–574. doi:10.1126/science.1151915

[2] Salas, J.D. & Obeysekera, J. (2014). Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. *Journal of Hydrologic Engineering*, 19(3), 554–568. doi:10.1061/(ASCE)HE.1943-5584.0000820

[3] Khaliq, M.N., Ouarda, T.B.M.J., Ondo, J.-C., Gachon, P. & Bobée, B. (2006). Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *Journal of Hydrology*, 329(3–4), 534–552. doi:10.1016/j.jhydrol.2006.03.004

[4] Montanari, A. & Koutsoyiannis, D. (2014). Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resources Research*, 50(12), 9748–9756. doi:10.1002/2014WR016092

[5] Villarini, G., Serinaldi, F., Smith, J.A. & Krajewski, W.F. (2009). On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resources Research*, 45, W08417. doi:10.1029/2008WR007645

[6] Rigby, R.A. & Stasinopoulos, D.M. (2005). Generalized additive models for location, scale and shape. *Applied Statistics*, 54(3), 507–554. doi:10.1111/j.1467-9876.2005.00510.x

[7] Poveda, G., Jaramillo, L. & Vallejo, L.F. (2014). Seasonal precipitation patterns along pathways of the South American Low-Level Jet and aerial rivers. *Water Resources Research*, 50(1), 98–118. doi:10.1002/2013WR014087

[8] Poveda, G. & Álvarez, D.M. (2012). El colapso de la hipótesis de estacionariedad por cambio y variabilidad climática: implicaciones para el diseño hidrológico en ingeniería. *Revista de Ingeniería* (Universidad de los Andes, Bogotá), núm. 36, enero–junio 2012, pp. 65–76. ISSN 0121-4993. https://www.redalyc.org/articulo.oa?id=121025826012

[9] IDEAM (2013). *Zonificación y codificación de unidades hidrográficas e hidrogeológicas de Colombia*. Bogotá, Colombia: Instituto de Hidrología, Meteorología y Estudios Ambientales.

[10] IDEAM (2022). *Estudio Nacional del Agua 2022*. Bogotá, Colombia: Instituto de Hidrología, Meteorología y Estudios Ambientales.

[11] Fundación Universidad del Valle (2025). *Informe hidrológico, hidráulico y de inundabilidad. Proyecto Granja Solar Las Marías, Bolívar (Cauca)*. Technical report IDMAR-PV-CIV-IF-104-REV-A_R1, Edition 02, January 2025.

[12] INVIAS (2009). *Manual de drenaje para carreteras*. Bogotá, Colombia: Instituto Nacional de Vías.

[13] Burnham, K.P. & Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd edn.). Springer, New York. doi:10.1007/b97636

[14] Mann, H.B. (1945). Nonparametric tests against trend. *Econometrica*, 13(3), 245–259. doi:10.2307/1907187

[15] Kendall, M.G. (1975). *Rank Correlation Methods* (4th edn.). Charles Griffin, London.

[16] Sen, P.K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, 63(324), 1379–1389. doi:10.1080/01621459.1968.10480934

[17] Pettitt, A.N. (1979). A non-parametric approach to the change-point problem. *Applied Statistics*, 28(2), 126–135. doi:10.2307/2346729

[18] Sneyers, R. (1990). *On the Statistical Analysis of Series of Observations*. WMO Technical Note No. 143. World Meteorological Organization, Geneva.

[19] White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817–838. doi:10.2307/1912934

[20] Dickey, D.A. & Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366), 427–431. doi:10.1080/01621459.1979.10482531

[21] Torrence, C. & Compo, G.P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79(1), 61–78. doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2

[22] Filliben, J.J. (1975). The probability plot correlation coefficient test for normality. *Technometrics*, 17(1), 111–117. doi:10.1080/00401706.1975.10489279

[23] Cox, D.R. & Snell, E.J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society, Series B*, 30(2), 248–265. doi:10.1111/j.2517-6161.1968.tb00724.x

[24] Ljung, G.M. & Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303. doi:10.1093/biomet/65.2.297

[25] Hamed, K.H. & Rao, A.R. (1998). A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology*, 204(1–4), 182–196. doi:10.1016/S0022-1694(97)00125-X

[26] Katz, R.W., Parlange, M.B. & Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, 25(8–12), 1287–1304. doi:10.1016/S0309-1708(02)00056-8

[27] Buishand, T.A. (1982). Some methods for testing the homogeneity of rainfall records. *Journal of Hydrology*, 58(1–2), 11–27. doi:10.1016/0022-1694(82)90066-X

[28] Caporello, G. & Maravall, A. (2004). *Program TSW: Revised Reference Manual*. Documentos Ocasionales N.° 0408. Banco de España, Madrid. https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosOcasionales/04/Fic/do0408e.pdf

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

[30] R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

[31] 52North Initiative (2014). *ILWIS Open: Integrated Land and Water Information System* (Version 3.8). https://www.itc.nl/ilwis/

A R packages used

Table 10: R packages used in CFE-Colombia v1.0.0 (R 4.3.1, Windows 10 x64).

Package	Version	Main function
`gamlss`	5.5-0	Non-stationary GAMLSS models
`gamlss.dist`	6.1-1	GAMLSS distributions
`extRemes`	2.2-1	GEV fitting, parametric bootstrap
`evd`	2.3-7.1	Gumbel functions (evd::dgumbel)
`fitdistrplus`	1.2-4	Log-Normal and Normal MLE fitting
`Kendall`	2.2.2	Mann-Kendall test
`trend`	1.1.6	Sen slope, Pettitt test
`WaveletComp`	1.2	Morlet wavelet analysis
`lmtest`	0.9-40	Breusch-Pagan test (White)
`tseries`	0.10-58	ADF test
`moments`	0.14.1	Statistical moments
`ggplot2`	4.0.0	Visualisation
`gridExtra`	2.3	Figure panels
`dplyr`	1.1.4	Data manipulation
`readxl`	1.4.5	Excel data reading

B Complete basin morphometric parameters

The complete morphometric parameters of the Guachicono River basin, determined from a digital terrain model with 2.5 m resolution using the ILWIS GIS, are presented in Table 1 in Section 2.1. The hydrological closure point (PCH 1) is located at Magna-Sirgas West coordinates: E 1 007 470.089 m, N 715 797.984 m.