hydrology

MDPI

Article

A Hybrid Index-Flood and Non-Stationary Bivariate Logistic Extreme-Value Framework for Flood Quantile Estimation in Data-Scarce Mexican Catchments

Laura Berbesi-Prieto and Carlos Escalante-Sandoval *

Faculty of Engineering, National Autonomous University of Mexico, Cd. Universitaria, Coyoacan CDMX 04510, Mexico; lauraberbesi@comunidad.unam.mx
* Correspondence: caes@unam.mx

Abstract

Regional flood frequency analysis (RFFA) is a cornerstone for estimating design floods at ungauged or data-scarce sites by pooling information within hydrologically homogeneous regions. This study proposes and evaluates a hybrid RFFA framework that integrates the Index-Flood (IF) technique with a bivariate logistic extreme-value model whose marginal distributions are formulated under both stationary and non-stationary assumptions. Non-stationarity is incorporated through a covariate-dependent location parameter, using time and large-scale climate indices—the Pacific Decadal Oscillation (PDO) and the Southern Oscillation Index (SOI)—as explanatory variables. The proposed approach is applied to two contrasting hydrological regions in Mexico—RH10 (Sinaloa) and RH23 (Chiapas Coast)—to assess its performance under differing climatic and hydrological regimes. Model adequacy and stability are evaluated using likelihood-based goodness-of-fit criteria (log-likelihood and Akaike Information Criterion) and a leave-one-out (jackknife) cross-validation scheme embedded within the IF regionalization workflow. Results indicate that non-stationary bivariate formulations dominate model selection at most stations and yield stable regional growth curves, providing robust and engineering-relevant performance under cross-validation. Overall, the proposed framework offers a conservative and operational pathway for regional flood quantile estimation that bridges local data scarcity and regional hydrological characterization in environments influenced by climate variability and long-term change.

Keywords: regional flood frequency analysis; index flood; bivariate extreme values; logistic model; maximum likelihood estimation; non-stationarity; climate covariates

Check for updates

Academic Editor: Brunella Bonaccorso

Received: 26 January 2026
Revised: 22 February 2026
Accepted: 3 March 2026
Published: 5 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Reliable estimation of design floods is essential for the planning, design, and safety assessment of hydraulic infrastructure. In practice, however, sites of interest rarely coincide with active gauging stations. Moreover, hydrometric networks often contain discontinued stations or records that are short, fragmented, or incomplete. These limitations restrict the applicability of conventional at-site flood frequency analysis and motivate the use of Regional Flood Frequency Analysis (RFFA), whereby information from multiple stations within a hydrologically homogeneous region is pooled to reduce uncertainty in flood quantile estimates for specified return periods.

Among regional approaches, the Index-Flood (IF) and regression-based methods remain the most widely applied. The GREHYS intercomparison study [1] demonstrated

that IF and regression approaches tend to be practically equivalent in many regional estimation contexts and often yield higher design flood estimates than alternative pooling strategies. Despite their widespread use, a key limitation of conventional RFFA frameworks is their reliance on the assumption of stationarity, namely, that the probabilistic structure of flood records remains invariant over time.

Growing evidence indicates that this assumption is often violated due to land-use change, river regulation, and climate variability and change, which can alter both the magnitude and timing of hydrological extremes. O'Brien and Burn [2] showed that ignoring trends or shifts in flood behavior can lead to significant bias in design quantile estimates. Consequently, improving RFFA methodologies to explicitly account for non-stationary conditions has become an important research priority.

In response to this challenge, several studies have proposed adaptations of the Index-Flood framework to accommodate non-stationarity using a variety of regional strategies and modeling assumptions. For instance, O'Brien and Burn [2] investigated non-stationarity in Canadian streamflow by introducing time as a covariate in the location and scale parameters of the GEV, Generalized Logistic (GLO), and Generalized Normal (GNO) distributions, while ultimately retaining a stationary scale formulation. In contrast, Sung et al. [3] proposed a two-step approach for precipitation extremes in South Korea, whereby linear trends were first removed from each site's series before applying a conventional stationary regional frequency analysis based on L-moments to the residuals. Other research has emphasized comparative assessments of regional methods. Using simulated data, Nam et al. [4] evaluated alternative non-stationary regional approaches under the assumption of a non-stationary location parameter and stationary scale and shape parameters within the GEV framework. Similarly, Kalai et al. [5] compared the performance of several non-stationary Index-Flood techniques using Canadian streamflow data, focusing on models with time-varying location parameters. Collectively, these studies highlight two dominant strategies for incorporating non-stationarity in RFFA: (i) embedding non-stationarity directly within the parameters of the flood distribution, or (ii) applying stationary regional methods to detrended residual series.

Building upon these developments, the present study extends non-stationary regional frequency analysis in two keyways. First, non-stationarity is incorporated into the marginal distributions of a bivariate logistic extreme-value model, following the estimation strategy proposed by Berbesi-Prieto and Escalante-Sandoval [6]. Second, the resulting non-stationary bivariate at-site (or at-pair) quantile estimates are integrated into an Index-Flood regionalization framework. This approach allows for the evaluation of predictive performance through jackknife cross-validation and supports applications at ungauged sites via regression-based estimation of the index flood.

This study contributes to non-stationary RFFA in four ways: (1) it embeds covariate-dependent non-stationarity in the marginal distributions of a bivariate logistic extreme-value model, enabling dependence-aware estimation of design quantiles; (2) it exploits partially overlapping records via a maximum bivariate likelihood, improving data efficiency in fragmented networks; (3) it integrates the resulting BEV quantiles into a full Index-Flood regionalization and jackknife validation workflow; and (4) it demonstrates an end-to-end ungauged application by coupling the regional growth curve with regression-estimated index floods based on basin descriptors.

2. Materials and Methods

2.1. At-Site Modeling

This study employs two extreme value distributions: Gumbel (1) and Generalized Extreme Value (GEV) (2). These distributions are applied in both their stationary and non-

stationary forms. For the non-stationary models, the analysis focuses on a non-stationary location parameter, a common strategy adopted in related hydrological studies [2–5]. Their respective probability distribution functions are:

$$F(x) = e^{-e^{-\left(\frac{x-\mu(c)}{\sigma}\right)}} \tag{1}$$

$$F(x) = exp\left\{-\left[1+\xi\left(\frac{x-\mu(c)}{\sigma}\right)\right]^{-1/\xi}\right\},\ 1+\xi\left(\frac{x-\mu(c)}{\sigma}\right) > 0 \tag{2}$$

where μ, σ, ξ are the location, scale, and shape parameters, respectively.

Following the methodology applied by Berbesi-Prieto and Escalante-Sandoval [6], this study models non-stationarity in the location parameter, μ, as a linear function of a covariate, c, given by $\mu(c) = \mu_0 + \mu_1 c$. The stationary form is thus a nested case of this model where $\mu_1 = 0$, reducing the parameter to $\mu = \mu_0$. As utilized in their work, the covariates (c) chosen for this analysis are time, representing long-term gradual trends, and the large-scale climate indices PDO (Pacific Decadal Oscillation) and SOI (Southern Oscillation Index), which are used to represent interannual variability.

2.2. Bivariate Logistic Extreme-Value Model

Dependence between paired series is modeled using the bivariate logistic extreme-value formulation. Its corresponding probability cumulative and density distribution functions are:

$$H(x,y,\theta) = exp\left\{-\left[(-lnF(x))^m + (-lnF(y))^m\right]^{1/m}\right\}, \tag{3}$$

$$\begin{aligned} h(x,y,\theta) = {} & \tfrac{f(x)}{F(x)}\tfrac{f(y)}{F(y)}[-lnF(x)]^{m-1}[-lnF(y)]^{m-1} exp\left\{-\left[(-lnF(x))^m + (-lnF(y))^m\right]^{1/m}\right\} \\ & \left[(-lnF(x))^m + (-lnF(y))^m\right]^{\frac{1}{m}-2}\left\{(m-1) + \left[(-lnF(x))^m + (-lnF(y))^m\right]^{1/m}\right\} \end{aligned} \tag{4}$$

where

x, y = annual maximum discharge at two sites (or paired records);
$F(x)$, $F(y)$ = marginals CDF's;
θ = parameter vector;
m = bivariate association parameter $m \in [1, \infty)$.

2.2.1. Log-Likelihood Function

To exploit partially overlapping hydrometric records between two sites, parameter estimation is based on a maximum bivariate likelihood that combines (i) a univariate likelihood for years in which only one series is available and (ii) a bivariate likelihood for the common period. Consider a station pair with annual maxima $\{x_i\}$ and $\{y_i\}$. Let n_2 be the number of years in the overlap (common) period and let n_1 and n_3 be the numbers of years in the non-overlap periods at the beginning and end of the joint time span, respectively. For non-overlap years, denoted by p_i the available annual maxima in the first non-overlap segment and by q_i the available annual maxima in the second non-overlap segment; depending on the record alignment, p_i and/or q_i may correspond to either x or y.

Define indicator variables δ_1, δ_2, $\delta_3 \in \{0,1\}$ such that $\delta_j = 1$ when the corresponding segment exists $(n_j > 0)$ and $\delta_j = 0$ otherwise. The joint parameter vector θ includes the marginal parameters for $F(x)$ and $F(y)$ (stationary or non-stationary, depending on the

candidate model) and the dependence parameter m of the logistic model. Under this structure, the log-likelihood is:

$$lnL(\theta) = \delta_1 \left[\sum_{i=1}^{n_1} lnf(p_i \,|\theta) \right] + \delta_2 \left[\sum_{i=1}^{n_2} lnh(x_i, y_i, \,|\theta) \right] + \delta_3 \left[\sum_{i=1}^{n_3} lnf(q_i \,|\theta) \right] \quad (5)$$

where $h(x_i, y_i, \,|\theta)$ is the bivariate density of the logistic extreme-value model (Equation (4)), and $f(\cdot)$ is the appropriate univariate marginal density for the variable available in the non-overlap segment. For non-stationary candidates, the covariate value c_i (Time, PDO, or SOI) associated with year i is used to evaluate the covariate-dependent location parameter $\mu(c_i) = \mu_0 + \mu_1 c_i$ within $F(x)$ and/or $F(y)$.

All candidate models are fitted by Maximum Likelihood Estimation (MLE), i.e., by numerically maximizing Equation (5) to obtain $\hat{\theta}$ and the corresponding maximized log-likelihood $\updownarrow = lnL(\theta)$, subject to standard parameter constraints (e.g., $\sigma > 0$ for marginal scales and $m \geq 1$ for the dependence parameter).

2.2.2. Model Nomenclature and Selection

The marginal distributions $F(x)$ and $F(y)$ were modeled using two extreme value distributions: the Gumbel and the Generalized Extreme Value (GEV) distributions. Each distribution was implemented under both stationary and non-stationary assumptions. Accordingly, the resulting combinations of bivariate extreme value (BEV) models were systematically denoted as BEV11, BEV12, BEV21, and BEV22, where the indices 1 and 2 correspond to the Gumbel and GEV marginal distributions, respectively.

The parameters of all marginal distributions, in both stationary and non-stationary formulations, were estimated using the Maximum Likelihood Estimation (MLE) method. This approach involves the numerical maximization of the log-likelihood function to obtain the optimal set of parameters. Model selection for each station was then conducted by comparing the maximized log-likelihood ($\updownarrow$) values and the Akaike Information Criterion (AIC):

$$AIC \; = 2k - 2\updownarrow \quad (6)$$

where k is the total number of free parameters in θ, including: (i) the marginal parameters for X and Y (plus any non-stationary slope terms μ_1 when included), and (ii) the logistic dependence parameter m. Models are preferentially selected by minimum AIC (primary criterion) with maximum ℓ used as a consistent goodness-of-fit check when AIC differences are small.

The AIC provides a relative measure of model performance by balancing goodness-of-fit, as reflected by ℓ, against model complexity, thereby penalizing excessive parameterization.

To further assess the suitability of non-stationary formulations, the deviance function D, as defined by Coles [7], was employed to quantify the improvement in model fit relative to the stationary case. The deviance is defined as $D = 2\{\updownarrow_1(M_1) - \updownarrow_0(M_0)\}$, in this equation $\updownarrow_0(M_0)$ and $\updownarrow_1(M_1)$ denote the maximized log-likelihoods of the stationary model M_0 and the non-stationary model M_1, respectively. The deviance statistic quantifies the improvement in model fit achieved by introducing additional parameters. Large values of D indicate that M_1 provides a statistically superior representation of the data, whereas small values suggest that the added complexity of M_1 does not yield a meaningful improvement over M_0.

This framework provides a formal hypothesis-testing criterion for model selection. The simpler model (M_0) is rejected in favor of the more complex M_1 at a significance level α if the deviation D exceeds a critical value, C_α. The critical value corresponds to the $(1 - \alpha)$ quantile of the chi-square distribution $\mathcal{X}_k^2$, where k is the difference in dimensionality (i.e., the number of additional parameters) between models M_1 and M_0 [7]. In this study, a

significance level of $\alpha = 0.05$ was adopted. Accordingly, for $k = 1$ degree of freedom, the critical value is $C_{0.05} = 3.84$, corresponding to the 0.95 quantile of the $\mathcal{X}_k^2$ distribution.

Operationally, for each station pair, we evaluate BEV candidates under S and NS forms (with candidate covariates Time, PDO, and SOI), estimate parameters by MLE (Equation (5)), and select the model that provides the best trade-off between fit and parsimony (minimum AIC), while using the deviance test to confirm whether the improvement of NS over its nested S counterpart is statistically supported. The selection of covariates (Time, PDO, and SOI) is based on their documented influence in the study regions [8]. In this study, the scale and shape parameters were kept constant, while a linear trend was incorporated into the location parameter, following established methodologies applied in similar research [2,4,6].

2.3. Regionalization via the Index-Flood Method

The Index-Flood (IF) method [9] is a foundational approach within Regional Flood Frequency Analysis (RFFA) and is among the most widely used and accepted regionalization techniques worldwide. Its early development and operational adoption in the United Kingdom [10], together with initial methodological support in the hydrological literature [11], established its relevance in flood risk studies. The versatility of the IF method is evidenced by its successful application to annual maximum flood flows across a broad range of hydroclimatic regions, including Italy [12], Canada [13], and the United Kingdom [14].

Over time, the IF framework has undergone substantial methodological advancements. One major enhancement involved its integration with the L-moments approach, enabling more robust and efficient parameter estimation; this integration has been widely applied in studies conducted in India [15], the United States [16], and Iran [17]. Subsequently, the traditional univariate formulation—primarily focused on peak discharges—was extended to bivariate contexts using copula functions. This extension allows for the joint modeling of flood characteristics such as peak flow and flood volume, thereby providing a more comprehensive representation of flood behavior [18–20]. More recently, and of relevance to this study, the IF method has been adapted to account for non-stationary conditions. Such developments have been explored in research addressing precipitation variability in Europe [21], non-stationary streamflow behavior in Canada [2,5], and precipitation extremes in South Korea [3].

The IF method is founded on the assumption that sites within a hydrologically homogeneous region share a common flood frequency structure. Under this assumption, the flood quantile $\hat{Q}_T$ at any site for a given return period T can be expressed as the product of two components: (i) a site-specific scaling factor, known as the index flood (IF), and (ii) a dimensionless regional growth curve, $\hat{q}_T$. The regional growth curve is estimated using pooled flood data from all sites within the region and is typically modeled using an extreme value distribution such as the Gumbel or the Generalized Extreme Value (GEV) distribution. This curve captures the shared frequency behavior across the region, while the index flood accounts for local hydrological magnitude.

2.4. Andrews Curves for Regional Delineation

A prerequisite for the application of the IF method is the identification of hydrologically homogeneous regions, within which the assumption of a common regional growth curve is valid. Consequently, regional delineation and homogeneity testing constitute critical preliminary steps in RFFA. Regions are typically defined by grouping catchments with similar physiographic, climatic, and hydrological characteristics, such as basin area, elevation, precipitation regime, and flow variability.

In this study, regional delineation was performed prior to parameter estimation and model selection. This step ensures consistency between the theoretical assumptions of the IF framework and the statistical properties of the pooled dataset, thereby strengthening the robustness of both stationary and non-stationary regional flood frequency analyses presented in the following sections.

Andrews curves provide a visualization of multivariate similarity by mapping each-p-dimensional descriptor vector $x = (x_i, \ldots, x_p)$ to a function [22]:

$$A(t) = \frac{x_1}{\sqrt{2}} + x_2 sin(t) + x_3 cos(t) + x_4 sin(2t) + x_5 cos(2t) + \ldots, \; t \in [-\pi, \pi] \quad (7)$$

where $x_1, x_2, \ldots$ represent attributes such as drainage area (A), mean annual precipitation (MAP), drainage density (DD), etc. Stations from a homogeneous region tend to produce Andrews curves that cluster, indicating similar hydrological behavior and flood-generating mechanisms.

2.5. Model Evaluation Metric: RMSE

To complement likelihood-based model adequacy measures used for at-site and bivariate model selection (log-likelihood and AIC), the predictive performance of the regionalization workflow was quantified using an error metric that is directly interpretable in discharge units. Specifically, we assessed the accuracy of regional design-flow estimates under each validation setting (all-sites regional fit, leave-one-out jackknife growth-curve validation, ungauged scaling using regression-estimated index flood, and regression jackknife) using the Root Mean Square Error (RMSE) computed across design return periods.

For each station i, RMSE is defined as:

$$RMSEi = \sqrt{\frac{1}{\tau}\sum_{\tau=1}^{\tau}\left[\hat{Q}_i(T_\tau) - Q_i(T_\tau)\right]^2} \quad (8)$$

where $Q_i(T_\tau)$ denotes the reference (at-site) design quantile at station i for return period T_τ and $\hat{Q}_i(T_\tau)$ is the corresponding estimate produced by the regional IF procedure (or its cross-validated/ungauged variants. In this study $T_\tau \in \{1.1, 2, 5, 10, 20, 50, 100\}$ years, so that $\tau = 7$. All return periods are equally weighted in the RMSE calculation, and RMSE values are reported in m^3/s.

2.6. Summary of the Methodological Approach

To provide a clear and integrated view of how the different methodological components are combined in this study, a schematic overview of the complete Regional Flood Frequency Analysis framework is presented in Figure 1. The figure summarizes the sequential steps adopted, from the delineation of hydrologically homogeneous regions and at-site distribution fitting—under both stationary and non-stationary assumptions—to regionalization using the Index-Flood approach and the final estimation of at-site flood quantiles across multiple return periods. This flowchart is intended to clarify the logical structure of the analysis and the interactions between regionalization, model selection, and non-stationary modeling components.

Regional Flood Frequency Analysis Methodology

↓

Delineation of Homogeneous Region (Andrews curves)

↓

At-site parameters

Distributions:
- Gumbel and GEV
- Bivariate Extreme Value: BEV11, BEV12, BEV21 and BEV22

Each distribution:
→ Stationary
→ Non-stationary (Covariates: Time, PDO, SOI)

↓

Model selection → log-likelihood and AIC

↓

Regional Flood Frequency Analysis: Index Flood
(Done twice → All sites and Jackknife)

↓

Estimation of flood quantiles: $\hat{Q}_{i,T}$ for
T = 1.1, 2, 5, 10, 20, 50, 100 years

↓

Normalization of the at-site data
$\frac{\hat{Q}_{i,T}}{IF}$

↓

Regional Median Value for each T

↓

Regional Equation

↓

Computation of at-site quantiles

Figure 1. Regional Flood Frequency Analysis framework adopted in this study.

3. Results

3.1. Study Regions, Data Sources and Regional Delineation

To evaluate the proposed framework under contrasting hydroclimatic conditions, two Mexican hydrological regions were analyzed: RH10 (Sinaloa) and RH23 (Chiapas Coast) (Figure 2).

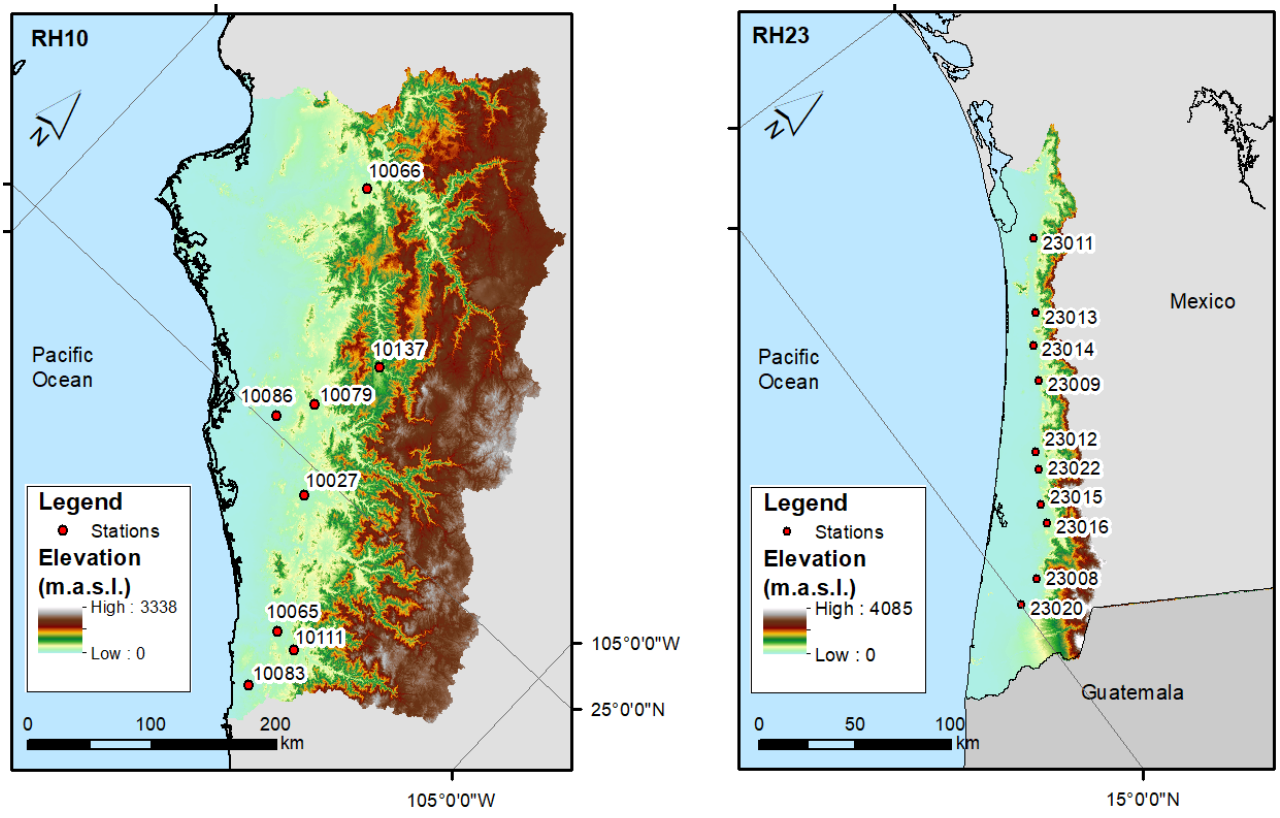


Figure 2. Geographic location of the two study regions in Mexico: Hydrological Region RH10 (Sinaloa) and Hydrological Region RH23 (Chiapas Coast). The figure provides the regional context of the analysis within the national territory; map scale and north arrow are included for reference.

RH10 (Sinaloa) is located in northwestern Mexico and comprises basins draining from the Sierra Madre Occidental toward the Pacific margin (Gulf of California). The region exhibits strong physiographic gradients—from steep headwaters to low-gradient coastal floodplains—and a pronounced wet-season regime. Extreme floods are frequently associated with tropical cyclone landfalls and monsoon-related convection, resulting in substantial interannual variability in annual maximum discharges.

RH23 (Chiapas Coast) occupies southeastern Mexico along the Pacific coastal corridor, spanning coastal plains and the adjacent Sierra Madre de Chiapas. Orographic enhancement, high atmospheric moisture availability, and intense storm systems generate frequent high-magnitude precipitation events and comparatively "flashy" runoff responses. Strong spatial heterogeneity across nearby basins makes RH23 a stringent test case for regionalization, particularly where classical assumptions of stationarity and regional homogeneity are more difficult to satisfy.

Annual maximum discharge series were obtained from the official hydrometric database maintained by the National Water Commission (CONAGUA) [23] (Appendix A). The spatial distribution of gauging stations in RH10 and RH23 is shown in Figure 3.

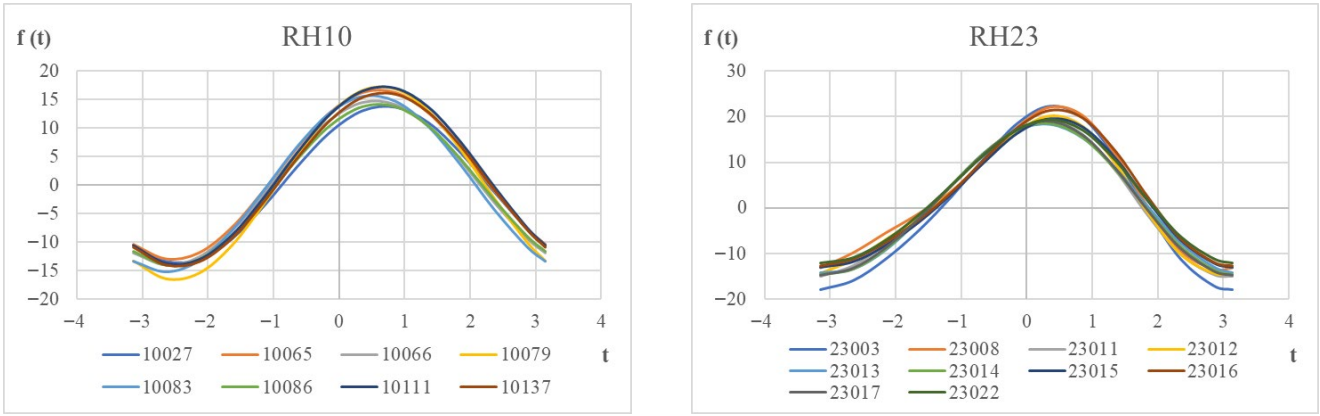


(**a**) (**b**)

Figure 3. DEM-based topographic (elevation) maps and gauging-station locations for the study regions: (**a**) RH10 (Sinaloa) and (**b**) RH23 (Chiapas Coast). Elevation is shown in m a.s.l. (orographic relief background), and stations used in the regional flood frequency analysis are indicated by red points and labels. Scale bars, north arrows, and elevation legends are included to improve cartographic interpretation.

Pooling groups were delineated using the Andrews curves method (Figure 4). The clustering inputs included morphometric and climatological descriptors (Appendix B) and their empirical association with return levels. The resulting groupings provide the working definition of "regional homogeneity" used in subsequent Index-Flood (IF) steps. Figure 4a

summarizes the RH10 delineation (adapted from [6]) and Figure 4b presents the RH23 outcome derived in this study.



(**a**) (**b**)

Figure 4. Andrews curves constructed from morphometric and climatological attributes. Clustering supports hydrological homogeneity of the study region. (**a**) RH10 (A, MAP, DD)—adapted from [6]; (**b**) RH23 (Em, MAP, DD, A).

3.2. Independence, Homogeneity, and Trend Diagnostics

Prior to frequency analysis and Index-Flood regionalization, the annual maxima series were screened for independence, homogeneity, and trend behavior (Tables 1 and 2). A symbolic notation is used: ✓ indicates the test condition is satisfied (e.g., independence is not rejected), whereas X indicates rejection of the null hypothesis.

Table 1. Results of the independence, homogeneity, and trend tests for the stations in Region RH10.

Station	I1	I2	I3	H1	H2	H3	H4	H5	H6	H7	H8	H9	T1
10027	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✗	✗
10065	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
10066	✓	✓	✓	✗	✓	✓	✗	✓	✓	✗	✗	✗	✗
10079	✓	✓	✓	✗	✓	✓	✗	✓	✓	✓	✓	✗	✓
10083	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✗
10086	✗	✓	✓	✗	✗	✓	✗	✗	✗	✗	✗	✗	✓
10111	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗
10137	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✗

Notes: (I) Independency: I1: Anderson; I2: Mantilla; I3: Amigó. (H) Homogeneity: H1: Helmert; H2: T-Student; H3: Cramer; H4: Pettit; H5: Normal Standard; H6: Buishand; H7: Von Neuman; H8: Thom; H9: Doorembos. (T) Tendency: T1: Mann–Kendall.

Table 2. Results of the independence, homogeneity, and trend tests for the stations in Region RH23.

Station	I1	I2	I3	H1	H2	H3	H4	H5	H6	H7	H8	H9	T1
23008	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
23009	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
23011	✓	✓	✓	✗	✗	✓	✗	✗	✗	✗	✓	✗	✓
23012	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗	✗
23013	✓	✓	✓	✓	✓	✓	✗	✗	✗	✓	✓	✓	✓
23014	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
23015	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
23016	✓	✓	✓	✗	✓	✓	✗	✗	✗	✗	✗	✗	✓
23020	✓	✓	✓	✗	✓	✗	✗	✗	✗	✗	✓	✗	✓
23022	✓	✓	✓	✗	✓	✓	✗	✗	✓	✗	✗	✗	✓

Notes: (I) Independency: I1: Anderson; I2: Mantilla; I3: Amigó. (H) Homogeneity: H1: Helmert; H2: T-Student; H3: Cramer; H4: Pettit; H5: Normal Standard; H6: Buishand; H7: Von Neuman; H8: Thom; H9: Doorembos. (T) Tendency: T1: Mann-Kendall.

RH10 (Table 1). All stations satisfy the independence assumption under the Anderson, Mantilla, and Amigó tests, supporting the use of annual maxima for extreme-value modeling. Homogeneity diagnostics indicate that most stations are broadly consistent with a homogeneous regional behavior under the multi-test ensemble; however, stations 10066 and 10086 show systematic inconsistencies, exhibiting multiple rejections across the homogeneity suite. Trend behavior based on Mann–Kendall indicates that strong monotonic trends are not widespread in RH10, suggesting that this region approximates a near-stationary regime for several sites, although localized non-stationary signals remain possible.

RH23 (Table 2). In contrast, RH23 exhibits markedly stronger departures from stationarity and homogeneity. While independence is generally supported for many stations, homogeneity tests show widespread rejections and Mann–Kendall results indicate significant trend behavior in most records. This diagnostic contrast is important for interpretation: RH10 represents a setting where classical regional assumptions are more defensible, whereas RH23 represents a more challenging environment in which non-stationary modeling and conservative regionalization are expected to be more critical.

3.3. At-Site Frequency Analysis

Following the methodology established by Berbesi-Prieto & Escalante-Sandoval [6], to identify the most statistically supported model at each observation site, the selection was based on two widely accepted criteria: the maximum log-likelihood value and the minimum Akaike Information Criterion (AIC). The identified best-fitting distributions for each gauging station in this study, along with their corresponding quantile estimates for return periods T = 1.1, 2 5, 10, 20, 50 and 100 years (Table 3).

The BEV21 configuration is selected for all RH10 stations, with the primary difference across sites being the selected covariate (Time or SOI). This uniformity suggests a stable regional dependence structure and supports the notion that a single bivariate formulation can represent RH10 behavior effectively (consistent with the more favorable screening outcomes in Table 1).

RH23 displays a more mixed selection across BEV classes (BEV11, BEV12, BEV21, BEV22), yet non-stationary covariates (notably Time and SOI, and in one case PDO) still dominate. The exception is station 23012, for which a stationary covariate entry is reported in Table 3. This station also appears among the comparatively more stable records in Table 2, providing a coherent explanation for why a stationary form may remain competitive there.

Overall, Table 3 indicates that incorporating non-stationarity in the marginal location parameter provides systematic gains in statistical support across most stations, which is consistent with the broader non-stationary frequency analysis literature that emphasizes covariate-dependent location as a parsimonious and effective structure [2–5].

3.4. Regional Index-Flood Results Using Observed Index Flood

The Index-Flood (IF) regionalization was then implemented following Figure 1. For each station, at-site quantiles (Table 3) were normalized by the site's index flood (mean annual maximum discharge) to obtain dimensionless growth factors. Regional growth values were computed as regional medians for each return period, and at-site quantiles were reconstructed by scaling the regional growth curve by each station's index flood.

To evaluate the performance of the regionalization, the Root Mean Square Error (RMSE) was calculated between these regional results and the at-site estimates previously listed in Table 3. The regional median values, their respective estimated quantiles, and the associated RMSE for regions RH10 and RH23 are presented in Table 4 and Table 5, respectively.

Table 3. Best-fitting probability distributions and estimated quantiles for various return periods at the study gauging stations.

Region	Station	Best-Fitting Distribution	Covariate	T (Years)						
				1.1	2	5	10	20	50	100
RH10	10027	BEV 21-10065	SOI	80.1	213.1	370.5	498.5	642.7	866.6	1067.0
	10065	BEV 21-10027	Time	390.8	908.3	1500.8	1969.8	2487.4	3272.2	3958.7
	10066	BEV 21-10079	SOI	106.0	227.9	375.1	496.8	635.6	854.1	1052.6
	10079	BEV 21-10066	Time	217.7	519.4	1004.1	1504.0	2185.0	3497.2	4943.4
	10083	BEV 21-10065	Time	93.9	268.1	540.6	815.8	1184.5	1881.3	2635.4
	10086	BEV 21-10137	Time	125.2	229.7	340.0	422.0	508.0	631.3	733.4
	10111	BEV 21-10065	Time	238.7	616.3	1228.3	1864.0	2734.8	4423.6	6296.2
	10137	BEV 21-10066	SOI	316.8	697.4	1343.0	2037.4	3016.0	4976.1	7217.6
RH23	23008	BEV21-23009	SOI	46.0	183.0	318.1	413.3	509.0	639.9	743.3
	23009	BEV11-23011	SOI	49.6	172.0	283.9	357.9	429.0	520.9	589.8
	23011	BEV21-23013	Time	7.9	59.4	136.3	211.2	308.6	486.7	673.4
	23012	BEV21-23016	-	106.6	252.6	337.6	376.9	405.4	432.4	447.3
	23013	BEV21-23015	Time	3.2	33.1	58.1	73.5	87.6	104.6	116.6
	23014	BEV22-23012	PDO	108.0	252.6	347.6	395.8	433.5	472.4	495.6
	23015	BEV22-23008	Time	23.2	92.1	173.8	240.2	315.0	431.3	535.4
	23016	BEV22-23012	Time	19.8	88.8	184.1	271.6	380.1	568.0	755.0
	23020	BEV21-23015	Time	90.1	194.6	291.4	356.2	418.9	500.8	562.8
	23022	BEV12-23012	Time	33.1	82.4	127.4	157.2	185.8	222.8	250.5

Note: Results for region RH10 are adapted from Berbesi-Prieto & Escalante-Sandoval [6], while results for region RH23 are original findings from the present study.

Table 4. Regional design flows (m^3/s) estimated using the Index-Flood technique and RMSE for Region RH10.

T (Years)	Regional Median	10027	10065	10066	10079	10083	10086	10111	10137
1.1	0.236	67.56	288.10	73.73	240.39	108.42	55.87	311.07	250.14
2	0.878	250.83	1069.63	273.74	892.50	402.53	207.42	1154.94	928.70
5	1.532	437.78	1866.86	477.76	1557.71	702.54	362.01	2015.74	1620.88
10	1.941	554.74	2365.62	605.40	1973.88	890.24	458.73	2554.29	2053.93
20	2.310	660.16	2815.17	720.45	2348.99	1059.42	545.90	3039.69	2444.25
50	2.756	787.55	3358.38	859.46	2802.24	1263.84	651.24	3626.22	2915.89
100	3.069	877.23	3740.83	957.34	3121.35	1407.77	725.40	4039.17	3247.94
RMSE (m^3/s)		86.05	264.45	77.27	801.47	528.36	35.93	1015.24	1709.85

Table 5. Regional design flows (m^3/s) estimated using the Index-Flood technique and RMSE for Region RH23.

T (Years)	Regional Median	23008	23009	23011	23012	23013	23014	23015	23016	23020	23022
1.1	0.185	49.80	40.63	23.39	47.01	10.40	37.22	31.66	30.55	67.63	22.14
2	0.693	186.54	152.18	87.64	176.10	38.98	139.41	118.61	114.42	253.35	82.92
5	1.215	326.81	266.61	153.53	308.51	68.28	244.24	207.80	200.46	443.85	145.26
10	1.542	414.79	338.38	194.86	391.56	86.67	309.99	263.74	254.43	563.33	184.37
20	1.837	494.20	403.17	232.17	466.52	103.26	369.33	314.23	303.13	671.18	219.67
50	2.194	590.26	481.54	277.30	557.21	123.33	441.13	375.31	362.06	801.65	262.37
100	2.446	657.96	536.76	309.10	621.12	137.47	491.72	418.36	403.59	893.59	292.46
RMSE (m^3/s)		37.94	29.72	162.14	92.57	14.11	76.51	52.54	157.26	218.34	28.39

In RH10 (Table 4), the regional median growth factors increase monotonically from 0.236 (T = 1.1 years) to 3.069 (T = 100 years), indicating a stable growth structure. RMSE values show that regional estimates generally track at-site quantiles well, with the smallest discrepancies at stations such as 10066 and 10086, and larger errors at high-magnitude stations (notably 10111 and 10137). This behavior is typical for IF procedures, which tend to smooth local extremes toward a regional central tendency.

In RH23 (Table 5), the regional median growth factors likewise increase consistently (0.185 at T = 1.1 years to 2.446 at T = 100 years). RMSE values are generally higher and more variable than in RH10, consistent with the stronger non-stationary and non-homogeneous signals indicated by Table 2. Stations 23009, 23013, and 23022 exhibit comparatively small RMSE, whereas stations such as 23011, 23016, and 23020 show larger discrepancies, reflecting a stronger influence of localized behavior and trends that are more difficult to represent with a single regional curve.

3.5. Jackknife (Leave-One-Out) Validation of the Regional Growth Curve

To evaluate predictive stability under an ungauged-like scenario, a leave-one-out jackknife was performed in which each station was excluded from the regional growth calibration and then predicted using the remaining stations (Tables 6 and 7). In this validation, the local index flood is assumed known (standard practice for isolating growth-curve performance).

Table 6. Regional design flows (m^3/s) estimated using the Index-Flood cross-validation (jackknife) procedure and RMSE for Region RH10.

T (Years)	10027	10065	10066	10079	10083	10086	10111	10137
1.1	66.55	269.58	69.18	257.62	115.94	52.16	333.38	252.86
2	252.31	1062.27	272.72	904.38	404.41	204.23	1170.30	910.41
5	444.28	1901.46	488.27	1541.33	687.36	364.58	1994.55	1568.40
10	565.11	2435.50	625.46	1933.48	861.23	466.44	2502.01	1976.44
20	674.36	2921.31	750.26	2283.87	1016.44	559.00	2955.44	2342.48
50	806.75	3513.01	902.28	2704.11	1202.43	671.64	3499.25	2782.96
100	900.17	3932.27	1010.00	2998.21	1332.51	751.40	3879.82	3092.07
RMSE (m^3/s)	79.30	307.69	84.54	848.96	564.35	43.29	1072.52	1788.35

Table 7. Regional design flows (m^3/s) estimated using the Index-Flood cross-validation (jackknife) procedure and RMSE for Region RH23.

T (Years)	23008	23009	23011	23012	23013	23014	23015	23016	23020	23022
1.1	54.26	35.32	26.81	40.33	11.76	31.93	36.28	34.86	61.03	19.07
2	190.07	141.73	92.21	166.12	41.54	130.68	125.60	119.84	246.81	78.58
5	323.65	255.65	155.80	303.00	70.97	237.70	212.80	202.43	446.68	143.38
10	405.84	328.52	194.72	391.22	89.13	306.55	266.27	252.97	574.83	185.14
20	479.26	394.99	229.39	472.02	105.36	369.54	313.95	297.98	691.87	223.40
50	567.29	476.15	270.85	571.02	124.84	446.66	371.02	351.81	834.92	270.27
100	628.88	533.77	299.81	641.50	138.49	501.53	410.90	389.40	936.55	303.65
RMSE (m^3/s)	52.724	36.01	166.7	103.3	15.84	80.61	56.86	164.3	239.4	33.23

RH10 (Table 6 vs. Table 4): Jackknife RMSE values increase modestly relative to the all-sites regional fit, with no evidence of instability or collapse of the regional curve. The increases are consistent with the expected loss of information when the target site is removed. This outcome provides empirical support that RH10 behaves as a coherent pooling group for IF regionalization.

RH23 (Table 7 vs. Table 5): Jackknife RMSE values are again higher and more variable than in RH10. Some stations (e.g., 23020) show larger deviations when excluded, indicating

sensitivity to local behavior that is not fully captured by the remaining stations. Importantly, even in these cases the jackknife outputs remain within an engineering-useful order of magnitude and tend to be conservative—an attribute that is desirable for design when direct data are limited.

3.6. Ungauged-Site Quantiles Using Regression-Estimated Index-Flood

Because ungauged sites lack observed discharge records, the index flood must be estimated from basin descriptors. Accordingly, the index flood for regions RH10 and RH23 ($\widehat{IF}_{RH10}$ and $\widehat{IF}_{RH23}$) were derived using a regional regression analysis that correlates mean annual maximum flow with selected physiographic and climatic descriptors [24]. Regression models were fit using the stations within each region with complete descriptor data. We used a log-linear formulation $ln(IF) = a_0 + a_1 ln(L) + a_2 ln(MAP) + \cdots$ which implies a multiplicative (power-law) model in original units and guarantee positive predictions, and the optimal configuration was selected based on the highest coefficient of determination (R^2). Regression sample sizes were n = 8 stations for RH10 and n = 10 stations for RH23 (all stations with complete descriptor data). In addition to R^2, model significance was evaluated using the overall F-test and coefficient-level t-tests.

For region RH10, drainage area initially exhibited a significant univariate correlation (R^2 = 0.74); however, predictive performance improved substantially when a multivariate approach was adopted. By integrating both morphometric and climatic variables, specifically main channel length (L) and mean annual precipitation (MAP), the model achieved an R^2 of 0.90. This result highlights the importance of interactions between basin morphology and climate in achieving accurate hydrological characterization. The resulting regional equation for RH10 is defined as follows:

$$\widehat{IF}_{RH10} = exp(0.79556 \cdot ln(L) + 2.6632 \cdot ln(MAP) - 15.3947) \quad (9)$$

Consistent with the findings for RH10, the index flood estimation for Region RH23 achieved its optimal fit through a multivariate linear function. This model, which incorporates two specific physiographic descriptors (main channel length (L) and mean basin elevation (E), provided superior explanatory power compared to simpler univariate alternatives. The mathematical expression defining this regional behavior for RH23 is:

$$\widehat{IF}_{RH23} = exp(-1.2671 \cdot ln(E) + 1.6483 \cdot ln(L) + 8.0381) \quad (10)$$

It is imperative to highlight that the explanatory variables (covariates) within this equation differ from those identified for RH10. This distinction underscores the methodological necessity of evaluating and testing diverse combinations of physiographic and climatic descriptor pairs to accurately capture the unique hydrological drivers specific to each region.

Following this framework, the index flood ($\widehat{IF}$) was estimated for each station in RH10—treated as a synthetic ungauged site—using Equation (9). These estimated scaling factors were then multiplied by the dimensionless regional growth factors presented in Table 4 to derive at-site quantile estimates for the various design return periods. The resulting estimates, along with the corresponding RMSE to assess predictive accuracy, are summarized in Table 8.

Table 8. Estimated index flood ($\widehat{IF}$), regional design flows (m^3/s) for ungauged catchments, and associated RMSE for Region RH10.

$\widehat{IF}$	**366.17**	**1160.18**	**423.52**	**781.90**	**321.30**	**233.74**	**1432.46**	**1101.71**
T (Years)	**10027**	**10065**	**10066**	**10079**	**10083**	**10086**	**10111**	**10137**
1.1	86.56	274.26	100.12	184.83	75.95	55.25	338.62	260.43
2	321.37	1018.24	371.71	686.24	281.99	205.14	1257.21	966.92
5	560.90	1777.17	648.75	1197.71	492.17	358.04	2194.24	1687.60
10	710.75	2251.97	822.07	1517.70	623.67	453.70	2780.48	2138.47
20	845.82	2679.93	978.30	1806.12	742.18	539.92	3308.87	2544.85
50	1009.03	3197.04	1167.06	2154.62	885.39	644.10	3947.34	3035.90
100	1123.93	3561.11	1299.97	2399.99	986.22	717.45	4396.85	3381.63
RMSE (m^3/s)	150.18	233.73	261.33	1100.74	750.92	34.36	953.02	1643.42

RH10 (Table 8 vs. Table 4). Replacing observed index flood with regression estimates increases uncertainty, and RMSE generally increases for several stations—especially where the regression error compounds with local growth-curve mismatch (e.g., station 10079). However, some stations remain stable (e.g., 10086), indicating that the regression is sufficiently informative for certain physiographic contexts.

A similar procedure was applied to region RH23, where the index flood ($\widehat{IF}$) was estimated using Equation (10). These values were subsequently scaled by the regional growth factors established in Table 5 to compute the design quantiles for the corresponding return periods. The resulting estimates and the associated predictive performance, as measured by the RMSE, are presented in Table 9.

Table 9. Estimated index flood ($\widehat{IF}$), regional design flows (m^3/s) for ungauged catchments, and associated RMSE for Region RH23.

$\widehat{IF}$	**267.62**	**202.15**	**179.49**	**263.42**	**58.38**	**141.60**	**178.61**	**152.37**	**350.87**	**130.80**
T (Years)	**23008**	**23009**	**23011**	**23012**	**23013**	**23014**	**23015**	**23016**	**23020**	**23022**
1.1	49.54	37.42	33.22	48.76	10.81	26.21	33.06	28.20	64.95	24.21
2	185.56	140.17	124.46	182.66	40.48	98.19	123.85	105.65	243.29	90.70
5	325.10	245.57	218.04	320.00	70.92	172.02	216.98	185.10	426.23	158.90
10	412.62	311.68	276.74	406.15	90.01	218.33	275.39	234.93	540.98	201.67
20	491.61	371.35	329.72	483.90	107.25	260.13	328.11	279.90	644.54	240.28
50	587.17	443.54	393.81	577.97	128.09	310.69	391.90	334.31	769.83	286.99
100	654.52	494.41	438.98	644.26	142.78	346.33	436.84	372.66	858.13	319.91
RMSE (m^3/s)	39.71	57.53	106.84	103.88	17.61	156.46	47.28	174.23	194.88	46.31

RH23 (Table 9 vs. Table 5). Error responses are mixed: several stations exhibit increased RMSE due to regression uncertainty (e.g., 23014), while other hydrologically complex stations show improved RMSE (e.g., 23011 and 23020). This pattern suggests that, for some records with strong non-stationary noise, regression-based index flood estimates can dampen the influence of local anomalies and produce more regionally consistent scaling.

3.7. Jackknife Validation of Ungauged Regression (Index-Flood Cross-Validation)

Finally, the regression equations for index flood were themselves evaluated under jackknife cross-validation (Tables 10 and 11), which more closely represents a practical ungauged application: the target station is excluded from regression calibration, and its index flood is predicted solely from basin descriptors.

Table 10. Leave-one-out cross-validated index flood ($\widehat{IF}$), regional design flows (m^3/s) for ungauged catchments, and associated RMSE for Region RH10.

$\widehat{IF}$	416.07	1114.83	870.46	256.95	251.17	231.79	1501.96	1117.02
T (Years)	**10027**	**10065**	**10066**	**10079**	**10083**	**10086**	**10111**	**10137**
1.1	96.89	246.60	193.07	65.10	63.50	51.16	380.50	266.92
2	367.32	971.71	761.14	228.52	221.47	200.31	1335.75	961.06
5	646.81	1739.36	1362.71	389.46	376.43	357.57	2276.53	1655.64
10	822.71	2227.88	1745.58	488.55	471.65	457.47	2855.72	2086.39
20	981.77	2672.27	2093.90	577.08	556.65	548.25	3373.25	2472.79
50	1174.50	3213.53	2518.17	683.27	658.50	658.73	3993.94	2937.78
100	1310.50	3597.04	2818.80	757.58	729.74	736.95	4428.32	3264.08
RMSE (m^3/s)	260.39	212.70	1244.55	2054.31	899.97	38.36	978.43	1700.94

RH10 (Table 10 vs. Table 8). As expected, RMSE values generally increase relative to the baseline regression case, reflecting the additional uncertainty of leaving the target site out of the regression fit. The magnitude of degradation varies by station, highlighting that regression performance is not uniform across physiographic regimes. Nonetheless, the method remains operationally feasible for design screening.

Table 11. Leave-one-out cross-validated index flood ($\widehat{IF}$), regional design flows (m^3/s) for ungauged catchments, and associated RMSE for Region RH23.

$\widehat{IF}$	266.83	197.93	200.46	265.92	62.93	128.48	179.56	146.33
T (Years)	**23008**	**23009**	**23011**	**23012**	**23013**	**23014**	**23015**	**23016**
1.1	53.81	31.86	42.52	42.23	13.17	20.40	38.09	30.92
2	188.52	127.82	146.26	173.95	46.50	83.50	131.85	106.27
5	321.01	230.56	247.12	317.28	79.46	151.90	223.39	179.50
10	402.53	296.27	308.85	409.65	99.78	195.89	279.52	224.31
20	475.35	356.22	363.84	494.26	117.95	236.15	329.56	264.23
50	562.66	429.41	429.61	597.92	139.76	285.43	389.47	311.96
100	623.75	481.38	475.53	671.72	155.04	320.49	431.33	345.29
RMSE (m^3/s)	55.57	70.03	104.24	118.00	26.88	176.92	51.55	188.78

RH23 (Table 11 vs. Table 9). Error increases are again present but remain within an engineering-interpretable range. Stations with stronger trend/homogeneity issues continue to show higher uncertainty; however, the overall stability of the predictions supports the utility of combining (i) a robust regional growth curve and (ii) basin-descriptor scaling when direct hydrometric information is absent.

3.8. Comparative Analysis: Non-Stationary vs. Stationary Models

To quantify the practical benefit of the proposed non-stationary framework relative to a conventional stationary approach, we conducted a benchmark comparison in RH10 using a stationary Generalized Extreme Value (GEV) model within the same Index-Flood (IF) and jackknife workflow. The stationary GEV is widely used as a baseline model in flood frequency analysis and provides a direct reference for assessing the value added by the non-stationary bivariate extreme-value (BEV) formulation.

The comparison was performed for stations 10027, 10065, 10066, and 10086, which present the smallest discrepancies under the IF regionalization in RH10 (Table 4) and therefore provide a stable baseline for isolating methodological differences. For each target station, a leave-one-out (jackknife) procedure was applied: the station was excluded from the regional growth-curve calibration, and design flows were predicted using the remaining stations.

Table 12 reports the resulting design flows obtained from the stationary GEV–IF jackknife procedure. Model performance was quantified using the RMSE, computed against

the corresponding at-site (best-fitting) quantiles reported in Table 3, consistent with the validation strategy used for the BEV results. Comparing Table 12 with the non-stationary jackknife results in Table 6 shows a systematic degradation in predictive accuracy under the stationary framework. Specifically, RMSE increases from 79.3 to 210.9 m^3/s at station 10027, from 307.7 to 1211.8 m^3/s at 10065, from 84.5 to 301.6 m^3/s at 10066, and from 43.3 to 228.8 m^3/s at 10086. These increases indicate that, even in RH10—where diagnostic screening suggests comparatively weak non-stationary signals—the non-stationary BEV-based regionalization yields substantially more accurate regional predictions than a stationary GEV baseline.

This result supports the use of covariate-dependent marginals within a dependence-aware BEV structure as a practically meaningful improvement over a stationary regional baseline.

RH10 was selected as the primary benchmark region because its stronger homogeneity and weaker trend diagnostics provide a more controlled baseline for isolating the incremental value of the non-stationary BEV structure relative to a stationary univariate GEV–IF workflow.

Table 12. Regional design flows (m^3/s) estimated using the stationary GEV Index-Flood cross-validation (jackknife) procedure and RMSE for selected stations in Region RH10.

T (Years)	1.1	2	5	10	20	50	100	RMSE (m^3/s)
10027	53.1	274.1	550.3	740.1	920.1	1147.3	1313.1	210.9
10065	227.1	1177.9	2369.6	3189.6	3967.9	4951.4	5669.0	1211.8
10066	58.1	301.5	606.4	816.3	1015.5	1267.1	1450.8	301.6
10086	44.0	228.1	458.8	617.5	768.2	958.6	1097.5	228.8

4. Discussion

4.1. Evaluation of Statistical Assumptions and Regional Stationarity

In region RH10, independence is consistently supported by the Anderson, Mantilla, and Amigó tests (Table 1). Homogeneity screening indicates that most stations are suitable for regional analysis under the adopted multi-test rule, while a small subset exhibits systematic inconsistencies and is therefore treated cautiously. Trend diagnostics further indicate that RH10 is largely stationary, with significant trends confined to a minority of sites.

Region RH23 exhibits markedly different behavior. Both homogeneity and trend tests indicate widespread departures from stationarity (Table 2). This contrast is informative, as it provides a natural stress test for the proposed framework: RH10 approximates a "classical" regionalization environment, whereas RH23 represents a more complex setting in which stationarity-based methods are expected to degrade.

Although the proposed framework focuses on statistical non-stationarity represented by time and large-scale climate indices (PDO and SOI), regional flood behavior can also be influenced by non-climatic drivers that alter runoff generation and flow routing. Land-use/land-cover (LULC) change, wetland modification, and river/floodplain engineering can modify infiltration capacity, storage, and connectivity, thereby affecting both the magnitude and variability of annual flood maxima. Because these factors were not explicitly parameterized in the present modeling (e.g., as covariates or as regionalization descriptors beyond the morphometric set), they provide important context for interpreting differences between RH10 and RH23 and are relevant to the residual heterogeneity observed in the diagnostic screening and IF performance.

Over approximately the last five decades, both regions have undergone substantial land-use/land-cover (LULC) change, although along contrasting trajectories shaped by climate, physiography, and dominant production systems [25–27]. In RH10, the prevailing

pathway has been the consolidation of a highly managed agro-hydrological mosaic, characterized by the expansion and intensification of irrigated agriculture in coastal valleys and lowlands, growth of urban and infrastructure corridors, and sustained pasture/rangeland pressure in foothills and accessible uplands. In the Sierra Madre Occidental headwaters, forest cover change has been more heterogeneous, reflecting disturbance/fragmentation in some areas and localized recovery in others, driven by land access, fire, and management conditions [26,27]. In RH23, LULC change has been dominated by conversion of tropical forest, riparian vegetation, and floodplain mosaics—particularly across the coastal plain and lower foothills—toward cropland and pasture, together with expansion and intensification of perennial commercial systems. On the Sierra Madre de Chiapas slopes, change often reflects reconfiguration of coffee-based agroecosystems across shade–sun management gradients, alongside localized deforestation for agriculture and grazing and patches of secondary vegetation where land has been abandoned or protected [26,27]. In both regions, coastal wetlands and mangroves have experienced fragmentation and hydrologic alteration linked to drainage works, channelization, settlement expansion, and land conversion; these changes are especially consequential where lagoon–estuary connectivity and floodplain storage historically moderated extremes [28–30].

From a hydrologic standpoint, a consistent consequence of these LULC shifts is a tendency toward greater runoff efficiency and disturbed sediment regimes, although the dominant controls differ by region [31–33]. Conversion of forest and riparian cover to agriculture and pasture—and soil compaction under grazing or intensive management—can reduce interception and infiltration, promote flashier hydrographs and higher event runoff coefficients, and increase erosion and sediment yield, with downstream impacts that include channel aggradation, reservoir/lagoon siltation, and water-quality degradation [31,33]. In RH10, flood regimes are additionally shaped by regulation and water use (e.g., irrigation withdrawals/returns, drainage networks, and floodplain modifications), which can modify seasonality and baseflow behavior independently of climate forcing. In RH23, the combination of steep orography, humid-tropical rainfall, and reduced floodplain/wetland connectivity increases sensitivity to rapid runoff production and reduced natural buffering, potentially intensifying compound flood processes [29,30]. Overall, both regions have shifted from more continuous native cover toward more fragmented, production- and infrastructure-dominated landscapes, implying that contemporary flood variability reflects not only climatic drivers but also the cumulative imprint of land management and coastal-plain modification [25,26,28].

These land-surface and connectivity changes provide a plausible complementary explanation for why RH23 exhibits stronger departures from homogeneity and stationarity diagnostics, and why dependence-aware non-stationary modeling can be more consequential in that setting.

Soil texture, depth, and hydraulic conductivity can also influence infiltration-excess and saturation-excess runoff generation and therefore affect regional flood response; these factors were not explicitly included in the present regionalization descriptors and should be incorporated in future extensions.

4.2. Performance of Non-Stationary Bivariate Modeling

Results in Table 3 show that non-stationary bivariate formulations dominate model selection for most stations across both regions, with only limited exceptions. This outcome is consistent with the broader non-stationary extreme-value literature (e.g., [2–5]), which frequently demonstrates that allowing distribution parameters—particularly the location parameter—to vary with time or climate indices improves model fit when trends or climate-driven shifts are present.

Relative to the predominant approach in prior studies, which typically relies on univariate GEV-type models with time-varying parameters (e.g., [2,4,5]), the primary contribution here lies in the use of a bivariate logistic extreme-value dependence structure combined with non-stationary marginals. This formulation provides a mechanism to (i) exploit paired information, particularly under partially overlapping records through the likelihood structure (Equation (4)), and (ii) reduce sensitivity to localized noise through three coupled mechanisms: (1) dependence-aware pooling—when sites share regional flood-generating mechanisms, the logistic dependence structure regularizes extremes by constraining joint tail behavior; (2) efficient use of partially overlapping records—the maximum bivariate likelihood (Equation (5)) uses univariate segments and the common bivariate segment simultaneously, reducing parameter uncertainty relative to discarding non-overlap years; and (3) shrinkage toward coherent regional behavior—under strong inter-site dependence, site-specific anomalies are less influential on fitted tail behavior.

4.3. Index-Flood Robustness Under Regionalization and Cross-Validation

Moving toward implementation in ungauged contexts, analysis of predictive errors (RMSE) relative to the global regional adjustment (Table 4) for region RH10 reveals a heterogeneous but expected response. For example, station 10066 exhibits a pronounced increase in RMSE, rising from 77.27 to 261.33 m^3/s, reflecting compounded uncertainty associated with regression-based index-flood estimation. In contrast, stations 10065 and 10086 display stable or slightly reduced error metrics. This behavior indicates that the regional framework—combining regression-derived index-flood estimates with the regional growth curve—captures the magnitude of extreme events with accuracy comparable to approaches based on observed at-site means, thereby validating the applicability of Equation (9) for discharge prediction in ungauged catchments within this region.

A similar comparison between the ungauged scenario (Table 9) and the baseline regional adjustment (Table 5) for region RH23 reveals a more divergent error response. As anticipated, stations 23009 and 23014 experience increases in RMSE (from 29.72 to 57.53 m^3/s and from 76.51 to 156.46 m^3/s, respectively), reflecting uncertainty introduced by replacing observed discharge records with regression-derived basin descriptors. Of particular interest, however, is the behavior observed at hydrologically complex catchments such as stations 23011 and 23020. In these cases, the use of regression-estimated index floods substantially improves predictive performance, with RMSE decreasing from 162.14 to 106.84 m^3/s at station 23011 and from 218.34 to 194.88 m^3/s at station 23020. This suggests that, for highly heterogeneous sites, regional regression can attenuate the influence of local outliers or non-stationary noise in individual records, yielding a more stable regional representation.

Comparison of cross-validation results (Table 10) with baseline regression estimates (Table 8) reveals a systematic increase in RMSE across region RH10. This trend is an expected consequence of the leave-one-out (jackknife) procedure, which increases predictive uncertainty by excluding the target station from calibration and relying solely on physiographic and climatic descriptors. However, the degradation in performance is not uniform. Stations 10086 and 10111 exhibit remarkable stability, with negligible RMSE variation relative to baseline regression results, confirming the robustness of index-flood estimation under synthetic ungauged conditions.

Consistent patterns emerge in region RH23 when contrasting cross-validation results (Table 9) with the baseline regional adjustment (Table 11). Despite modest RMSE increases due to the exclusion of local station data—illustrated by station 23008, where RMSE rises from 39.71 to 55.57 m^3/s—errors remain within acceptable engineering limits. Notably, stations 23011 and 23015 show exceptional consistency, with RMSE values remaining nearly

invariant (106.84 to 104.24 m^3/s and 47.28 to 51.55 m^3/s, respectively). These results further confirm the robustness of the physiographic regression equations for estimating design discharges in ungauged catchments across both regions.

4.4. Limitations and Future Works

4.4.1. Limitations

Despite the methodological advances presented in this study, several limitations should be acknowledged.

1. Scope of regions and sample size.
   The analysis is restricted to two hydrological regions. Although these regions were selected to represent contrasting hydroclimatic conditions, a broader application across additional regions within Mexico would be required to assess the generality and robustness of the proposed framework.
2. Structure of non-stationarity.
   Non-stationarity is incorporated exclusively through a linear covariate effect on the location parameter. While this formulation is both common and theoretically defensible in hydrological frequency analysis [2–5], it may be insufficient in settings where changes in variability or tail behavior are significant, or where covariate effects exhibit nonlinear or threshold-driven dynamics.
3. Treatment of dependence structure.
   The dependence parameter of the logistic model (m) is assumed to be constant over time. Dependence strength between flood characteristics may vary in response to evolving climate regimes or long-term hydroclimatic shifts. Ignoring potential non-stationarity in dependence may limit model flexibility in highly dynamic environments.

4.4.2. Future Work

Several avenues for future research emerge from the limitations identified above.

1. Expanded covariate structures.
   Future studies could consider additional covariates and more flexible formulations, including nonlinear effects, interaction terms (e.g., PDO × SOI), or regime-switching models to better capture complex climate–flood relationships.
2. Non-stationarity beyond the location parameter.
   Allowing non-stationarity in the scale parameter—and potentially the shape parameter—could improve model realism in regions experiencing changes in flood variability or tail behavior. Deviance-based nested hypothesis testing could be employed to control over-parameterization.
3. Time-varying dependence modeling.
   Extending the framework to permit temporal or covariate-dependent variation in the dependence parameter (m) would provide valuable insight into the evolution of joint flood behavior under changing climate conditions.

5. Conclusions

This research developed and validated a hybrid methodological framework integrating the Index-Flood technique with non-stationary multivariate models. By implementing a robust multi-test ensemble for evaluating independence and homogeneity, the proposed approach significantly reduces uncertainty in quantile estimation for data-scarce basins. In region RH10, the non-stationary BEV21 distribution demonstrates high consistency and predictive stability under leave-one-out cross-validation. Although region RH23 exhibits greater statistical complexity and widespread non-stationarity, the framework effectively

captures these dynamics through the inclusion of non-stationary covariates such as time and the Southern Oscillation Index (SOI).

Validation metrics confirm that the Index-Flood technique provides a robust approximation of design discharges, with RMSE values generally representing a minor fraction of total flow. A key finding concerns regression-based estimation under ungauged conditions: while marginal error increases are observed at hydrologically stable sites, multivariate regression substantially improves predictive accuracy at more complex stations—such as 23011 and 23020—by mitigating the influence of local non-stationary noise and outliers. Consequently, the regional growth curve supported by non-stationary covariates is validated as an effective representation of homogeneous regional behavior. Overall, the framework yields conservative and reliable estimates, offering a robust tool for hydraulic engineering design and decision-making in non-stationary environments.

Application to regions RH10 (Sinaloa) and RH23 (Chiapas Coast) demonstrates that:

a. Non-stationary bivariate formulations are strongly supported by likelihood-based model selection for most stations, particularly within the more complex RH23 environment.
b. Index-Flood regionalization based on at-site and at-pair quantiles produces stable regional growth curves and maintains acceptable predictive performance under jackknife cross-validation, especially in RH10.
c. For ungauged applications, multivariate regression provides a practical pathway for estimating the index flood; although uncertainty increases when the index flood is inferred from basin descriptors, the resulting quantiles remain operationally useful and conservative for design screening.
d. Compared with conventional stationary regional flood frequency analysis (RFFA), the proposed framework explicitly incorporates climate and time covariates while preserving the interpretability and practical implementation of Index-Flood regionalization.

Overall, the proposed framework constitutes a practical and defensible tool for flood-risk assessment in data-scarce catchments under non-stationary conditions, supporting more reliable design quantiles when traditional stationarity assumptions are no longer valid.

Author Contributions: L.B.-P. and C.E.-S. both contributed to all aspects of the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data will be made available at a reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BEV	Bivariate Extreme Value Distribution
GEV	Generalized Extreme Value
IF	Index Flood
MAP	Mean Annual Precipitation
PDO	Pacific Decadal Oscillation
RFFA	Regional Flood Frequency Analysis
SOI	Southern Oscillation Index
T	Return Period

Appendix A

Inventory of gauging stations analyzed in RH10 and RH23. For each station, the table reports the period of annual-maximum discharge observations, record length (years), and geographic location (latitude and longitude in decimal degrees).

Station	Period	Length of Record (Years)	Latitude (°)	Longitude (°)
RH 10				
10027	1938–1995	58	24.8042	107.1458
10065	1953–1999	47	23.9556	106.5958
10066	1955–2005	51	26.7361	108.3292
10079	1959–1999	41	25.3333	107.5375
10083	1960–1992	33	23.5125	106.4833
10086	1960–1992	33	25.0931	107.6944
10111	1958–2009	52	23.9333	106.4250
10137	1958–2008	51	25.8556	107.3764
RH23				
23008	1959–1996	38	15.1333	92.4667
23009	1961–2005	45	15.7000	93.2167
23011	1961–2005	45	16.0708	93.7500
23012	1962–1991	30	15.4875	92.9500
23013	1964–2002	39	15.8667	93.4833
23014	1964–1997	34	15.7500	93.3333
23015	1964–2005	42	15.3500	92.7167
23016	1964–1998	35	15.3083	92.7333
23020	1964–1994	35	15.0000	92.4167
23022	1964–1992	29	15.4500	92.8958

Appendix B

Catchment physiographic and climatic characteristics for stations in RH10 and RH23. Variables include Area (A, km^2), MAP (mm), Perimeter (P, km), Main stream length (L, km), Mean Elevation (E, m.a.s.l.), Slope (S,%), and dimensionless morphometric indices (Compactness Coefficient—Kc, Basin Form Factor—Rf) together with drainage density DD. These descriptors were used as inputs to the Andrews-curves regionalization and to support index-flood estimation.

Station	A (km^2)	MAP (mm)	P (km)	L (km)	E (m.a.s.l.)	S (%)	Kc (-)	Rf (-)	DD
RH 10									
10027	388.26	912.90	122.60	52.01	297.78	0.14	1.79	0.33	0.21
10065	6102.18	871.50	725.79	258.89	1506.42	0.12	2.61	0.12	0.25
10066	1373.58	777.45	255.08	106.91	995.28	0.14	1.93	0.18	0.25
10079	1010.89	1031.19	208.96	89.76	868.91	0.18	1.85	0.21	0.27
10083	827.43	754.80	185.92	83.41	305.49	0.13	1.81	0.22	0.28
10086	226.59	831.20	96.59	40.49	227.62	0.10	1.60	0.34	0.24
10111	5277.52	993.50	646.25	217.63	1683.94	0.12	2.50	0.12	0.25
10137	3300.19	982.90	438.28	162.18	1754.24	0.06	2.01	0.12	0.24
RH23									
23008	354.58	2637.90	105.16	56.01	1298.17	0.22	1.57	0.30	1.46
23009	205.88	2213.50	68.78	29.75	711.30	0.20	1.35	0.33	1.45
23011	167.74	1701.70	68.13	26.43	669.83	0.16	1.48	0.41	1.49
23012	282.06	2285.00	85.16	38.52	807.73	0.18	1.43	0.30	1.45
23013	52.41	2340.80	35.21	14.58	749.64	0.28	1.37	0.67	1.43
23014	76.27	2139.30	45.28	23.49	692.78	0.20	1.46	0.59	1.46
23015	178.83	3287.90	67.13	35.13	973.61	0.24	1.42	0.38	1.38
23016	223.93	3551.70	98.85	40.07	1309.73	0.26	1.86	0.44	1.42
23020	321.44	3990.00	102.44	47.84	853.94	0.15	1.61	0.32	1.53
23022	124.90	2597.70	69.30	35.32	1253.75	0.22	1.75	0.55	1.35

References

1. Groupe de Recherche en Hydrologie Statistique (GREHYS). Inter-Comparison of Regional Flood Frequency Procedures for Canadian Rivers. *J. Hydrol.* **1996**, *186*, 85–103. [CrossRef]
2. O'Brien, N.L.; Burn, D.H. A Nonstationary Index-Flood Technique for Estimating Extreme Quantiles for Annual Maximum Streamflow. *J. Hydrol.* **2014**, *519*, 2040–2048. [CrossRef]
3. Sung, J.H.; Kim, Y.-O.; Jeon, J.-J. Application of Distribution-Free Nonstationary Regional Frequency Analysis Based on L-Moments. *Theor. Appl. Climatol.* **2018**, *133*, 1219–1233. [CrossRef]
4. Nam, W.; Kim, S.; Kim, H.; Joo, K.; Heo, J.-H. The Evaluation of Regional Frequency Analyses Methods for Nonstationary Data. *Proc. Int. Assoc. Hydrol. Sci.* **2015**, *371*, 95–98. [CrossRef]
5. Kalai, C.; Mondal, A.; Griffin, A.; Stewart, E. Comparison of Nonstationary Regional Flood Frequency Analysis Techniques Based on the Index-Flood Approach. *J. Hydrol. Eng.* **2020**, *25*, 06020003. [CrossRef]
6. Berbesi-Prieto, L.; Escalante-Sandoval, C. Flood Frequency Analysis Using the Bivariate Logistic Model with Non-Stationary Gumbel and GEV Marginals. *Hydrology* **2025**, *12*, 274. [CrossRef]
7. Coles, S. *An Introduction to Statistical Modeling of Extreme Values*; Springer Series in Statistics; Springer: London, UK, 2001.
8. Alvarez-Olguin, G.; Escalante-Sandoval, C. Modes of Variability of Annual and Seasonal Rainfall in Mexico. *JAWRA J. Am. Water Resour. Assoc.* **2017**, *53*, 144–157. [CrossRef]
9. Dalrymple, T. *Flood Frequency Analysis*; US Geological Survey Water-Supply Paper 1543-A; United States Government Printing Office: Washington, DC, USA, 1960.
10. Hosking, J.R.M.; Wallis, J.R.; Wood, E.F. An Appraisal of the Regional Flood Frequency Procedure in the UK *Flood Studies Report*. *Hydrol. Sci. J.* **1985**, *30*, 85–109. [CrossRef]
11. Cunnane, C. Methods and Merits of Regional Flood Frequency Analysis. *J. Hydrol.* **1988**, *100*, 269–290. [CrossRef]
12. Brath, A.; Castellarin, A.; Franchini, M.; Galeati, G. Estimating the Index Flood Using Indirect Methods. *Hydrol. Sci. J.* **2001**, *46*, 399–418. [CrossRef]
13. Grover, P.L.; Burn, D.H.; Cunderlik, J.M. A Comparison of Index Flood Estimation Procedures for Ungauged Catchments. *Can. J. Civ. Eng.* **2002**, *29*, 734–741. [CrossRef]
14. Kjeldsen, T.R.; Jones, D. Estimation of an Index Flood Using Data Transfer in the UK. *Hydrol. Sci. J.* **2007**, *52*, 86–98. [CrossRef]
15. Kumar, R.; Chatterjee, C. Regional Flood Frequency Analysis Using L-Moments for North Brahmaputra Region of India. *J. Hydrol. Eng.* **2005**, *10*, 1–7. [CrossRef]
16. Lima, C.H.R.; Lall, U.; Troy, T.; Devineni, N. A Hierarchical Bayesian GEV Model for Improving Local and Regional Flood Quantile Estimates. *J. Hydrol.* **2016**, *541*, 816–823. [CrossRef]
17. Malekinezhad, H.; Nachtnebel, H.P.; Klik, A. Comparing the Index-Flood and Multiple-Regression Methods Using L-Moments. *Phys. Chem. Earth Parts A/B/C* **2011**, *36*, 54–60. [CrossRef]
18. Chebana, F.; Ouarda, T.B.M.J. Index Flood–Based Multivariate Regional Frequency Analysis. *Water Resour. Res.* **2009**, *45*, 2008WR007490. [CrossRef]
19. Ben Aissia, M.-A.; Chebana, F.; Ouarda, T.B.M.J.; Bruneau, P.; Barbet, M. Bivariate Index-Flood Model: Case Study in Québec, Canada. *Hydrol. Sci. J.* **2015**, *60*, 247–268. [CrossRef]
20. Requena, A.I.; Chebana, F.; Mediero, L. A Complete Procedure for Multivariate Index-Flood Model Application. *J. Hydrol.* **2016**, *535*, 559–580. [CrossRef]
21. Hanel, M.; Buishand, T.A.; Ferro, C.A.T. A Nonstationary Index Flood Model for Precipitation Extremes in Transient Regional Climate Model Simulations. *J. Geophys. Res. Atmos.* **2009**, *114*, 2009JD011712. [CrossRef]
22. Andrews, D.F. Plots of High-Dimensional Data. *Biometrics* **1972**, *28*, 125–136. [CrossRef]
23. CONAGUA. Banco Nacional de Datos de Aguas Superficiales (BANDAS). Available online: https://app.conagua.gob.mx/bandas/ (accessed on 1 April 2024).
24. Rao, A.R.; Hamed, K.H. Regional Frequency Analysis of Wabash River Flood Data by L-Moments. *J. Hydrol. Eng.* **1997**, *2*, 169–179. [CrossRef]
25. CONAGUA. *Regiones Hidrológicas/Regionalización Hidrológica de México (Reference Material on Hydrological Regions, Including RH10 and Costa de Chiapas)*; CONAGUA: Mexico City, Mexico, 2018.
26. INEGI. Conjunto de Datos Vectoriales de Uso Del Suelo y Vegetación (Series I–VI) (National Land Use/Vegetation Maps). Available online: https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=702825572136 (accessed on 1 April 2024).
27. FAO. *Global Forest Resources Assessment 2020*; Food and Agriculture Organization of the United Nations: Rome, Italy, 2020.
28. CONABIO. *Monitoreo/Inventario de Manglares y Humedales de México*; Geospatial Datasets and Technical Reports; Comisión Nacional Para El Conocimiento y Uso de La Biodiversidad: Mexico City, Mexico, 2020.
29. Alongi, D.M. Mangrove Forests: Resilience, Protection from Tsunamis, and Responses to Global Climate Change. *Estuar. Coast. Shelf Sci.* **2008**, *76*, 1–13. [CrossRef]

30. Barbier, E.B.; Hacker, S.D.; Kennedy, C.; Koch, E.W.; Stier, A.C.; Silliman, B.R. The Value of Estuarine and Coastal Ecosystem Services. *Ecol. Monogr.* **2011**, *81*, 169–193. [CrossRef]
31. Bruijnzeel, L.A. Hydrological Functions of Tropical Forests: Not Seeing the Soil for the Trees? *Agric. Ecosyst. Environ.* **2004**, *104*, 185–228. [CrossRef]
32. Calder, I.R. Forests and Water—Ensuring Forest Benefits Outweigh Water Costs. *For. Ecol. Manag.* **2007**, *251*, 110–120. [CrossRef]
33. Foley, J.A.; DeFries, R.; Asner, G.P.; Barford, C.; Bonan, G.; Carpenter, S.R.; Chapin, F.S.; Coe, M.T.; Daily, G.C.; Gibbs, H.K.; et al. Global Consequences of Land Use. *Science* **2005**, *309*, 570–574. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.