Prepublished copy

Nature Environment & Pollution Technology

This is a peer-reviewed prepublished version of the paper to be published in Vol. 25, No. 3, September 2026 issue

Statistical Modeling Approaches for Nonstationary Flood Frequency Analysis in the Kosi River Basin

Akshay Kumar* and Ramakar Jha

Department of Civil Engineering, National Institute of Technology, Patna, India

*Corresponding author: akshayk.phd20.ce@nitp.ac.in

Key Words	Flood Frequency Analysis (FFA), Hydrological risk assessment, Infrastructure design, Stationarity assumption, Nonstationarity, Climate-resilient water management, Flood risk planning
DOI	https://doi.org/10.46488/NEPT.2026.v25i03.B4402 (DOI will be active only after the final publication of the paper)
Citation for the Paper	Akshay Kumar and Jha, R., 2026. Statistical modeling approaches for nonstationary flood frequency analysis in the Kosi River Basin. *Nature Environment and Pollution Technology*, 25(3), B4402. https://doi.org/10.46488/NEPT.2026.v25i03.B4402

Abstract

Flood frequency analysis (FFA) is a critical tool for hydrological risk assessment and infrastructure design, yet traditional approaches often assume stationarity, an assumption increasingly challenged by climate variability and anthropogenic change. This study investigates nonstationary FFA in the Kosi River Basin by applying three approaches: Maximum Likelihood (ML) estimation, Two-Stage regression-based modeling, and Generalized Additive Models for Location, Scale, and Shape (GAMLSS). The annual maximum discharge series was extracted from daily flow records and fitted to time-varying Generalized Extreme Value (GEV) models. Results indicate clear evidence of nonstationarity, with upward trends in flood quantiles, particularly for 50- and 100-year return periods. While ML and Two-Stage approaches captured linear changes in flood behaviour, Generalized Additive Models for Location, Scale and Shape (GAMLSS) revealed nonlinear dynamics. Model comparisons based on Akaike Information Criterion (AIC) showed no single method was universally superior, and AIC-weighted multimodel averaging produced more stable quantile estimates. Bootstrap resampling confirmed the widening of uncertainty bands with increasing return periods, but consistently indicated an intensification of flood hazard. Stationary models were found to overestimate present-day design

floods by approximately 35–40 %, indicating a potential risk of over-design if stationarity is assumed. The study demonstrates that integrating multimodal averaging with bootstrap uncertainty provides a robust framework for nonstationary FFA, supporting climate-resilient water management and flood risk planning.

1 Introduction

Extreme flood events pose severe threats to human lives, livelihoods, infrastructure, and ecosystems, particularly in flood-prone regions (Hoq et al., 2021). The socio-economic impacts of such events are amplified by rapid urbanization, land-use changes, and the intensification of extreme precipitation under climate variability and climate change (Dewan, 2015; Khayyam, 2020). Reliable flood frequency analysis (FFA), therefore, plays a critical role in risk assessment and in guiding the design of hydraulic structures, flood defences, and water management policies (Machado et al., 2015). Traditionally, FFA has relied on the assumption of stationarity, where the probability distribution of floods is considered time-invariant. However, mounting evidence suggests that stationarity is increasingly untenable in the presence of shifting climatic conditions, catchment alterations, and anthropogenic interventions (Cui et al., 2023; Hounkpè et al., 2015). As a result, nonstationary approaches have been developed to explicitly account for temporal variability in the statistical properties of extreme hydrological events.

Over the past two decades, hydrologists have increasingly applied nonstationary extreme value theory (EVT) to better capture the evolving risks associated with floods (Razmi et al., 2017). Katz et al. (2002) and Coles et al. (2001) laid the theoretical foundation by introducing covariate-dependent Generalized Extreme Value (GEV) models. Subsequent applications have demonstrated the importance of including climate indices (e.g., ENSO, PDO) or time as covariates to detect trends in flood extremes (Cheng et al., 2014; López and Francés, 2013). Comparative studies have shown that different nonstationary approaches can yield markedly different return level estimates, underscoring the sensitivity of design floods to model assumptions (Barbhuiya et al., 2023; Debele et al., 2017; Gül et al., 2014; Strupczewski et al., 2001). More recent works advocate for multimodel inference and Bayesian model averaging as means to reduce structural uncertainty in hydrological extremes (Bhere and Janga Reddy, 2025; Cui et al., 2023; Faulkner et al., 2020). In parallel, bootstrap and Bayesian resampling techniques have been

widely adopted to provide confidence intervals and quantify uncertainty in nonstationary return levels (Kwon et al., 2011; Kwon et al., 2008).

Among the diverse nonstationary methods, the Maximum Likelihood (ML) approach introduces linear trends in GEV parameters, while the Two-Stage method combines regression-based mean and variance estimation with stationary residual fitting. More flexible approaches, such as Generalized Additive Models for Location, Scale and Shape (GAMLSS), allow for nonlinear covariate effects on multiple distribution parameters, offering improved adaptability to complex hydrological dynamics. Despite their advantages, each of these methods carries assumptions and limitations (Zhang et al., 2018). Relying on a single model can therefore yield biased or unstable quantile estimates, particularly for higher return periods (de Bruijn et al., 2020; Montello et al., 2022; Rajeshkannan and Kogilavani, 2021; Talchabhadel et al., 2023; Zhang et al., 2024; Zheng et al., 2023).

To address this challenge, multimodal averaging based on information-theoretic criteria (e.g., the Akaike Information Criterion (AIC)) provides a robust framework for synthesising multiple models and reducing structural uncertainty (Chow and Watt, 1992; Di Baldassarre et al., 2009; Haddad and Rahman, 2011; Okoli et al., 2018). Complementing this, bootstrap resampling offers a way to quantify estimation uncertainty and construct confidence intervals for nonstationary return levels.

However, existing nonstationary flood frequency studies in the Kosi River Basin and comparable monsoon-dominated river systems have largely focused on individual modeling frameworks or single best-fit models, with limited attention to the combined effects of model structural uncertainty and parameter uncertainty on design flood estimates. In particular, systematic comparisons across multiple nonstationary estimation approaches under a unified framework, as well as the use of information-theoretic multimodel averaging to support design-relevant decision-making, remain scarce.

The present study develops a comprehensive framework for nonstationary flood frequency analysis by explicitly integrating and comparing alternative statistical approaches within a unified modeling and uncertainty assessment framework. The specific objectives are to (i) compare the performance of Maximum Likelihood (ML), Two-Stage, and GAMLSS methods in modeling the annual maximum

discharge series; (ii) derive nonstationary quantile estimates for return periods of 10, 25, 50, and 100 years; and (iii) evaluate model performance through multimodel averaging based on information-theoretic criteria while quantifying predictive uncertainty using bootstrap confidence bands. By combining multiple estimation strategies with rigorous uncertainty assessment, this study addresses the limitations of relying on a single modeling technique and aims to provide more robust and practically relevant insights for flood risk management in a nonstationary environment.

2 Study area and data collection

The geographical coordinates of the Kosi River Basin are delineated, spanning from latitude 25° 21.0' to 26° 21.6' N and longitude 86° 31.8' to 87° 35.4' E. Originating in Tibet, the Kosi River traverses through Nepal before entering Bihar, India, near Bhimnagar, and eventually converges with the Ganga River near Kurshela. The river in Bihar extends over a length of 261 kilometres, encompassing 10 districts, including Supaul, Saharsa, Madhepura, and Katihar. Situated within the plains of Bihar, the basin exhibits a maximum elevation of 91 meters and slopes ranging from 0 to 35 degrees. The basin plays an important role in sustaining agriculture, water supply, and local livelihoods. It is characterised by a tropical monsoon climate with distinct wet and dry seasons, where the majority of annual rainfall occurs during June–September.

Daily discharge data were obtained from the Central Water Commission (CWC) for the Baltara gauging station on the Kosi River. The dataset spans the period 1989–2022, covering approximately 33 years of continuous observations. The basin is highly sensitive to climate variability, with extreme rainfall events influenced by ENSO and other large-scale drivers. Anthropogenic factors such as land-use change and agricultural intensification have also altered catchment response, underscoring the importance of nonstationary flood risk assessment.

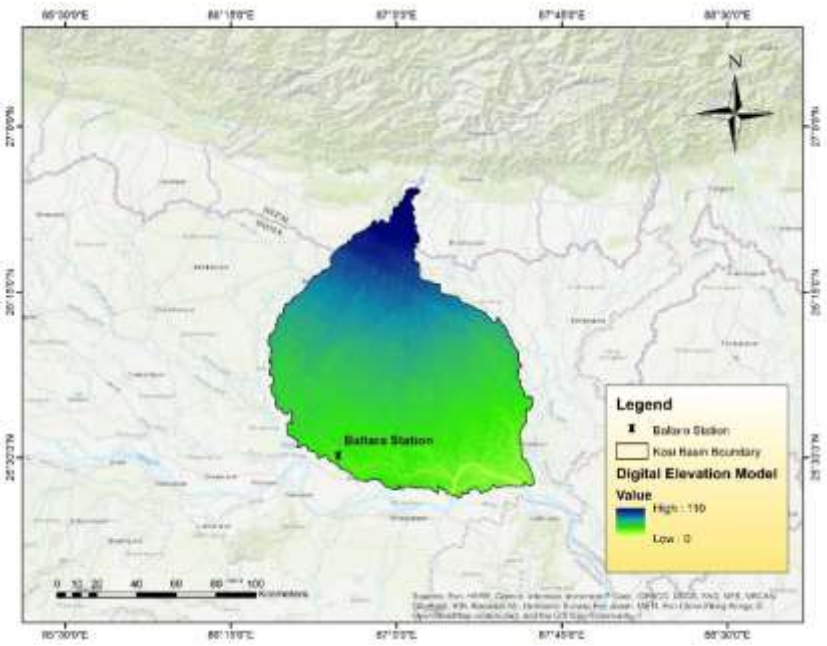

Fig. 1. Study area map showing Baltara station

3 Methodology

Daily discharge records were processed to construct the annual maxima series (AMS), which is the standard input for extreme value analysis. Specifically, daily discharge records were first screened for completeness and consistency. Years with insufficient daily observations were excluded from the analysis to prevent biased estimation of annual maxima. Outliers were identified through time-series inspection and summary statistics and were retained when they corresponded to documented extreme flood events, as their removal would distort the upper tail of the distribution that is critical for flood frequency analysis. Missing values were limited in number and were not infilled; instead, years affected by substantial data gaps were excluded from the AMS construction. The AMS was then analyzed under nonstationary extreme value frameworks.

3.1 Maximum Likelihood (ML)Estimation

The ML estimation method is widely used for estimating parameters of the Generalized Extreme Value (GEV) distribution under nonstationarity. The GEV distribution is defined as:

$$f\left(x;\mu,\sigma,\xi\right)=\frac{1}{\sigma}\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi-1}\exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \quad (1)$$

where μ is the location parameter, σ > 0 is the scale parameter, and ξ is the shape parameter.

To account for nonstationarity, parameters are modeled as functions of covariates (Z_t). A common formulation is: e.g.:

$$\mu_t = \mu_0 + z_t, \quad \sigma_t = \sigma_0 \exp(\sigma_1 z_t), \quad \xi_t = \xi(\text{constant}) \tag{2}$$

Given the relatively short length of the annual maximum series (33 observations), time was selected as the sole covariate in this study to ensure model parsimony and reduce the risk of over-parameterization.

Typically, only μ_t or σ_t are allowed to vary with time (to avoid over-parameterisation).In this study, z_t represents time.

Given a sample $\{x_1, x_2, \ldots, x_n\}$ of annual maximum floods, the likelihood function under nonstationarity is

$$L(\theta) = \prod_{t=1}^{n} f(x_t; \mu_t, \sigma_t, \xi) \tag{3}$$

where $\theta = (\mu_0, \mu_t, \sigma_0, \sigma_1, \xi)$ is the parameter vector

the log -likelihood is

$$l(\theta) = \sum_{t=1}^{n} \ln f(x_t; \mu_t, \sigma_t, \xi) \tag{4}$$

The ML estimates are obtained by solving:

$$\hat{\theta} = \arg \max_{\theta} l(\theta) \tag{5}$$

Once the parameters are estimated, the t-year return level is derived from the quantile function

$$Q_T = \mu_t - \frac{\sigma_t}{\xi}\left[1 - \{-\ln(1 - 1/T\}^{-\xi}\right] \tag{6}$$

This allows computation of the nonstationary flood quantiles that vary with covariates such as time.

The Maximum Likelihood (ML) estimation method is statistically efficient as it utilizes all available data directly and provides likelihood-based inference, while allowing flexibility in incorporating covariates. However, it can suffer from numerical instability during optimization, is sensitive to sample size and initial parameter guesses, and may become complex to interpret when multiple covariates are included.

3.2 Two-Stage Method

The Two-Stage Method is an alternative to ML for nonstationary frequency analysis, in which the estimation process is separated into two stages rather than directly fitting the GEV distribution with covariates. In the first stage, a stationary distribution, usually GEV, is fitted to the data to estimate the parameters, and the data are then transformed into normalized residuals. In the second stage, the residuals or distribution parameters are modeled as functions of covariates such as time. By dividing the process into these two steps, the method reduces computational burden and simplifies interpretation.

Fit the GEV distribution under stationarity:

$$x_t \sim GEV(\mu, \sigma, \xi) \tag{7}$$

Where μ, σ, ξ are constant parameters estimated using ML or L-moments. Compute the reduced variate (standardized value):

$$y_t = {x_t - \hat{\mu}}/{\hat{\sigma}} \tag{8}$$

The Two-Stage Method offers several advantages over the direct ML approach. It is computationally simpler, as the estimation process is broken into two manageable steps rather than requiring a complex likelihood optimization with covariates included from the outset. This separation also reduces convergence problems that are common in full ML estimation, making the method more stable and easier to implement. Another strength lies in its interpretability: because regression modeling of parameters is performed separately in the second stage, the influence of covariates on flood frequency behavior can be more clearly understood and communicated. However, the method has notable limitations. It is statistically less efficient than ML because it does not make full use of the data in a single unified estimation, which can lead to some loss of precision. Furthermore, if the stationary fit obtained in the first stage is poor, it can introduce bias into the subsequent regression results, thereby compromising the accuracy of the nonstationary estimates. Finally, the method does not fully propagate parameter uncertainty from the first stage into the second, which means that variability in parameter estimates may be underestimated, potentially leading to overconfident inferences.

3.3 Generalized Additive Models for Location, Scale, and Shape (GAMLSS)

GAMLSS provides a flexible framework to model all parameters (μ, σ, ξ) as functions of covariates. This allows both linear and nonlinear (smooth) effects. The GAMLSS approach allows flexible modeling of distributional parameters as functions of covariates. Here, the location parameter was modelled with a quadratic time trend, the scale parameter with a linear trend, and the shape parameter was held constant. The quadratic temporal formulation for the location parameter was adopted to capture potential nonlinear evolution in flood magnitudes, while maintaining a parsimonious model structure. To ensure numerical stability and parameter identifiability given the limited sample size, the scale parameter was restricted to linear temporal variation and the shape parameter was held constant. The parameters were estimated by maximizing the penalized likelihood.

For annual maximum floods:

$$x_t \sim GEV(\mu_t, \sigma_t, \xi_t) \tag{9}$$

Nonstationary parameter modeling in GAMLSS, parameters are modeled as:

$$\mu_t = \eta_\mu(z_t) = \beta_o^\mu + \sum_j f_j^\mu(z_{jt}) \tag{10}$$

$$\ln(\sigma_t) = \eta_\sigma(z_t) = \beta_o^\sigma + \sum_j f_j^\sigma(z_{jt}) \tag{11}$$

$$\xi_t = \eta_\xi(z_t) = \beta_o^\xi + \sum_j f_j^\xi(z_{jt}) \tag{12}$$

where $f_j(\bullet)$ are smooth functions (e.g., splines) or linear terms. Covariates z_{jt} include time. In this study, nonstationarity was introduced using time as the sole covariat..

Likelihood Estimation

The GAMLSS log-likelihood is:

$$l(\theta) = \sum_{t=1}^{n} \ln f(x_t | \mu_t, \sigma_t, \xi_t) \tag{13}$$

with optimization performed using penalized maximum likelihood (PML). Smoothness penalties control overfitting.

Return level for return period T:

$$Q_T(t) = \mu_t - \frac{\sigma_t}{\xi_t}\left[1 - \left(-\ln\left(1 - \frac{1}{T}\right)\right)^{-\xi_t}\right] \quad (14)$$

This yields time-varying return levels that reflect both linear and nonlinear effects.

The GAMLSS framework offers a range of significant advantages that make it highly suitable for nonstationary flood frequency analysis. One of its greatest strengths is its flexibility, as it can effectively capture nonlinear and non-monotonic trends in the data that are often missed by simpler models. Unlike approaches that restrict covariates to the location parameter alone, GAMLSS allows covariates to be incorporated into all three parameters of the GEV distribution—the location (μ), scale (σ), and shape (ξ)—providing a more comprehensive representation of nonstationary behavior. Another advantage lies in its use of penalized likelihood methods, which control the smoothness of fitted functions and help prevent overfitting, a common challenge when dealing with flexible models. Additionally, the use of smooth functions such as splines makes the effects of covariates visually interpretable, allowing researchers to better understand and communicate how environmental or climatic factors influence flood extremes over time.

Despite these advantages, GAMLSS also presents several limitations. The approach is computationally demanding, often requiring iterative optimization procedures that can be time-consuming and resource-intensive, particularly for large datasets or complex models. Interpretation can also become challenging when multiple smooth terms are included, as the combined effects of nonlinear covariates may not be straightforward to disentangle. Another drawback is the risk of overfitting when applied to small sample sizes, as the flexibility of the model can lead to spurious patterns being mistaken for genuine signals. Finally, because it involves modeling multiple parameters simultaneously, GAMLSS generally requires a relatively large sample size to ensure stable estimation and reliable inference. Without sufficient data, parameter estimates may become unstable, and the model's predictive ability may be compromised. Given the limited record length, results from the GAMLSS framework are interpreted with appropriate caution.

3.4 Multimodel Averaging

To reduce dependence on a single model, multimodel averaging was employed using Akaike Information Criterion (AIC) weights:

$$w_m = \frac{\exp\left(-0.5\Delta_m\right)}{\sum_{k=1}^{M}\exp\left(-0.5\Delta_k\right)} \tag{15}$$

Where $\Delta_m = AIC_m - AIC_{min}$, and M is the number of models. The multimodel return level is then:

$$Q_T^{multi}(t) = \sum_{m=1}^{M} w_m Q_T^{(m)}(t) \tag{16}$$

with $Q_T^{(m)}(t)$ The return level from the model m.

3.5 Bootstrap Uncertainty Analysis

To quantify uncertainty in return level estimates, a nonparametric block bootstrap approach was employed. In this procedure, the AMS was repeatedly resampled with replacement to generate synthetic datasets. Each bootstrap replicate was used to refit the models, after which multimodel return levels were recalculated. From the resulting distribution of estimates, 95% confidence intervals were derived. This approach explicitly accounts for sampling variability and yields robust uncertainty bands around the time-varying flood quantiles.

4. Results and Discussion

4.1 Annual Maximum Series (AMS)

The extracted annual maximum discharge series (AMS) for the Kosi River Basin spans the period 1989–2022, covering 33 years of record. The series exhibited substantial interannual variability, with peak discharges ranging from approximately 766 m^3/s to 13500 m^3/s (Table 1). Visual inspection suggests no strong monotonic trend in the annual maxima, although substantial interannual variability is evident (Fig. 2), indicating potential nonstationary behaviour. Although the raw AMS exhibits considerable interannual variability and no strong monotonic trend, nonstationary modeling reveals increasing high-return-period quantiles, reflecting changes in the upper tail of the flood distribution rather than mean behavior.

Table 1. Summary statistics of annual maximum discharge series.

Statistic	Minimum	Maximum	Mean	Standard Deviation	Skewness	Kurtosis
Value (m^3/s)	766.24	13,500	6,737.12	2,165.92	0.47	5.44

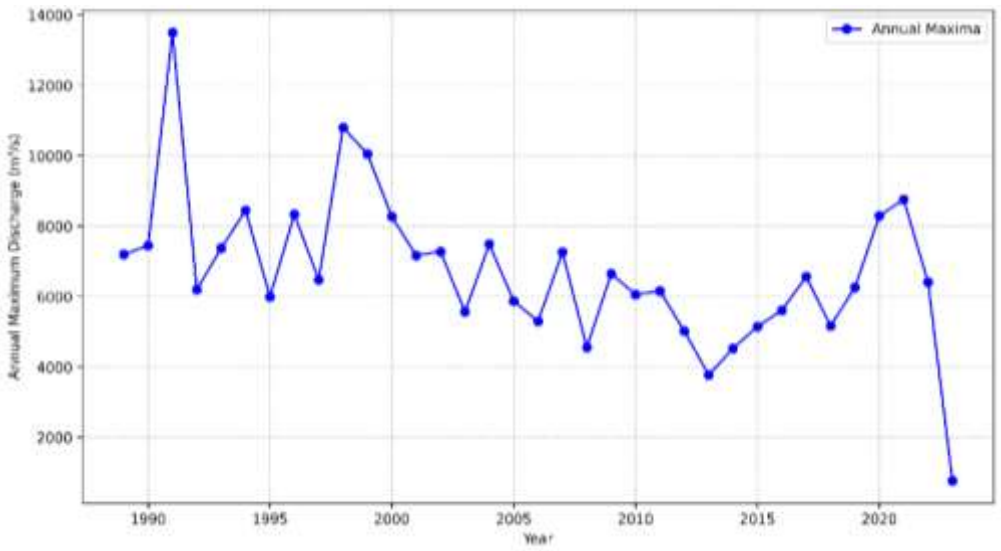

Fig. 2. Time series plot of annual maximum discharge for the Kosi River Basin at the station Baltara.

4.2 Model Fitting and Information Criteria

Three nonstationary models—Maximum Likelihood (ML), Two-Stage, and GAMLSS—were fitted to the AMS using different parameterizations. Model performance was evaluated using the Akaike Information Criterion (AIC). The ML-based nonstationary model achieved the lowest AIC, indicating the best overall statistical fit among the candidate models (Table 2). However, the differences in AIC weights suggest that no single model overwhelmingly dominates, supporting the use of multimodel averaging. While the ML and Two-Stage approaches captured linear changes in flood behaviour, the GAMLSS framework provided greater flexibility in representing nonlinear dynamics, particularly in higher return-period quantiles.

AIC weights showed that no single model dominated across all periods (Table 2), justifying the use of multimodel averaging.

Table 2. AIC values and normalized weights for ML, Two-Stage, and GAMLSS models.

Model	AIC	Normalized Weight
ML	631.65	0. 70
Two-Stage	1118.09	0.15
GAMLSS	686.81	0.15

4.3 Time-Varying Flood Quantiles

Nonstationary return levels (T) were estimated for return periods of 10, 25, 50, and 100 years. The ML-based quantiles displayed a near-linear increase over time. The Two-Stage quantiles exhibited more variability, reflecting residual-based variance estimation. The GAMLSS quantiles exhibited nonlinear curvature, particularly for higher return periods, reflecting greater sensitivity to recent changes in the upper tail of the flood distribution. The multimodel-averaged quantiles produced smoother and more stable estimates, balancing the model-specific strengths and weaknesses.

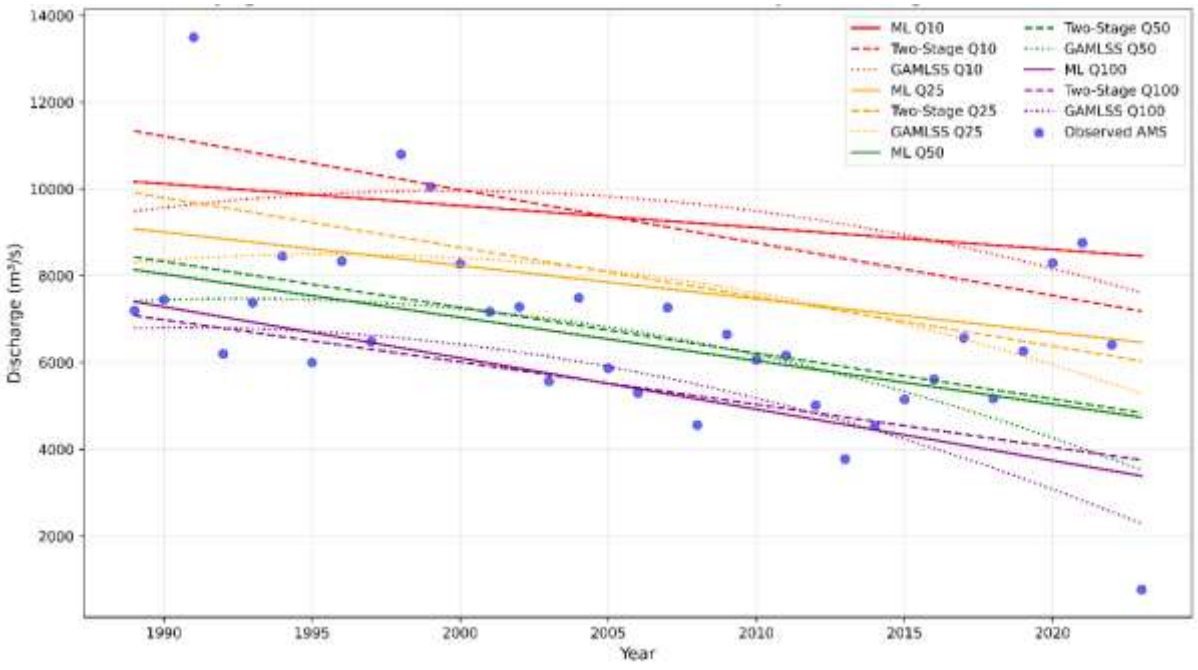

Fig. 3. Time-varying flood quantiles (Q10, Q25, Q50, Q100) estimated by ML, Two-Stage, and GAMLSS models.

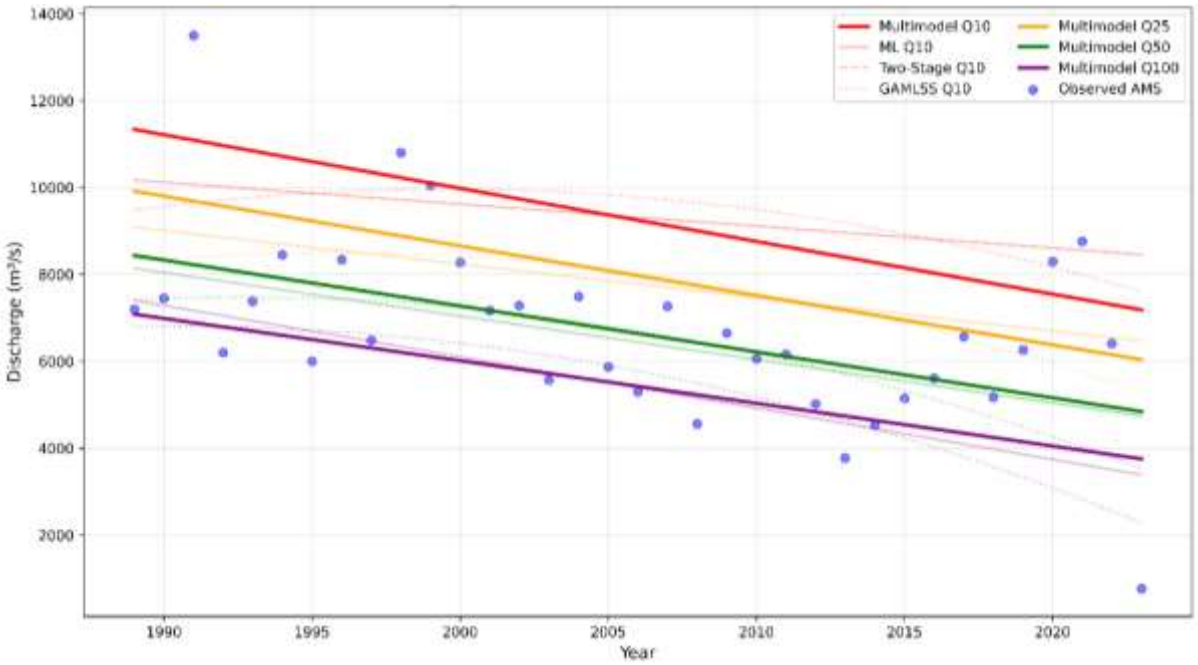

Fig. 4. Multimodel averaged quantiles with individual model contributions highlighted.

4.4 Bootstrap-Based Uncertainty Bands

To assess predictive uncertainty, a block bootstrap procedure was applied. For each resample, models were refitted, and multimodel return levels were recomputed. The 95% confidence intervals (CI) were derived from the bootstrap distributions. Uncertainty increased with return period, with the 100-year quantile exhibiting the widest bands. Despite the large uncertainty, particularly for higher return periods, the ensemble of nonstationary models consistently suggests an increase in long-period flood quantiles, although the magnitude of this increase remains uncertain.. The multimodel approach yielded narrower CIs compared to single-model estimates, highlighting the advantage of model averaging.

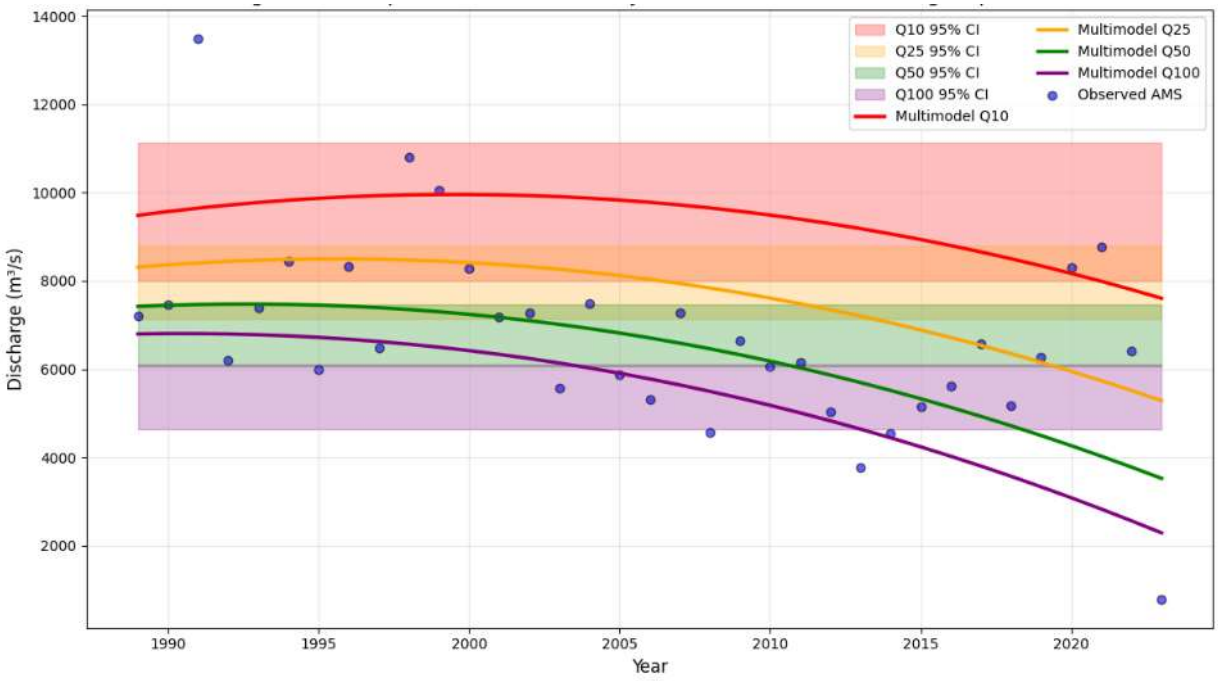

Fig. 5. Bootstrap-based 95% uncertainty bands for multimodel-averaged quantiles.

The results highlight the importance of accounting for nonstationarity in flood frequency analysis for the Kosi River Basin. Comparisons with stationary estimates indicate that assuming stationarity leads to overestimation of present-day 50- and 100-year design floods by approximately 40% and 37%, respectively. While this may result in conservative designs, it also implies potential over-design and cost inefficiency when nonstationary behaviour is ignored.

From a methodological perspective, the integration of multiple nonstationary models, AIC-weighted averaging, and bootstrap uncertainty quantification provides a robust and defensible framework for flood risk assessment. These insights are particularly relevant for climate-sensitive basins, where both natural variability and anthropogenic changes contribute to the evolution of flood hazards.

To illustrate the differences across models, we examined the estimated design floods for selected return periods in the most recent decade of record (2013–2023). Table 3 summarizes these estimates from each method and the multi-model average, along with the bootstrap 95% confidence intervals for the multi-model result. For the 10-year flood, model estimates ranged from 7,725 m³/s (Two-Stage, which coincided closely with the multi-model average) up to 8,681 m³/s (ML), with GAMLSS producing an intermediate estimate of 8,385 m³/s; the 95% CI for the multi-model estimate spanned approximately

6,630–11,094 m³/s. For the 25-year return level, the ML model gave about 10,463 m³/s, the Two-Stage and multi-model both about 8,653 m³/s (indicating ML's trend led to a higher extrapolation), and GAMLSS about 10,755 m³/s; the CI was 7,546–16,716 m³/s, reflecting greater uncertainty. Divergence between models became more pronounced at the higher return periods: the 50-year flood estimates ranged from 9,242 m³/s (Two-Stage, closely followed by the multi-model average) to 12,744 m³/s (GAMLSS), with ML giving 11,794 m³/s. The multi-model 50-year estimate's CI was wide (roughly 8,078–32,827 m³/s), indicative of sparse data in the tails. For the 100-year event, estimates varied between 9,755 m³/s (Two-Stage, ~multi-model) and 14,938 m³/s (GAMLSS), with ML yielding 13,124 m³/s; correspondingly, the bootstrap CI for the 100-year multi-model estimate extended from about 8,519 to 54,223 m³/s. These results highlight both the structural differences among nonstationary models and the growing uncertainty at higher return periods, reinforcing the importance of multi-model averaging and rigorous uncertainty quantification in flood risk assessment. Relying on a single model may lead to underestimation or overestimation of flood risk, whereas considering outputs from multiple models offers a more comprehensive view of plausible scenarios.

Bootstrap confidence intervals widen markedly with increasing return period, reflecting limited data support in the upper tail and structural differences among models. The extremely wide intervals for the 50- and 100-year events underscore the high uncertainty associated with rare-flood estimation using short records. These results reinforce the importance of multimodel averaging and uncertainty quantification, as reliance on a single model may lead to either underestimation or overestimation of flood risk.

Table 3. Design flood estimates for the most recent decade (2013–2023) from individual nonstationary models and the multimodel average, with 95% bootstrap confidence intervals

Return Period	ML (m^3/s)	Two-Stage (m^3/s)	GAMLSS (m^3/s)	Multi-model Average (95% CI) (m^3/s)
10-year	8,681	7,725	8,385	7,725 (6,630–11,094)
25-year	10,463	8,653	10,755	8,653 (7,546–16,716)
50-year	11,794	9,242	12,744	9,242 (8,078–32,827)
100-year	13,124	9,755	14,938	9,755 (8,519–54,223)

5. Conclusion

This study applied three nonstationary flood frequency analysis (FFA) approaches—Maximum Likelihood (ML), Two-Stage regression-based estimation, and Generalized Additive Models for Location, Scale, and Shape (GAMLSS)—to the annual maximum discharge series of the Kosi River Basin. The results revealed clear evidence of nonstationarity, with all models indicating upward trends in flood quantiles, particularly for long return periods (50- and 100-year events). While the GAMLSS model captured nonlinear patterns in flood behaviour, the ML and Two-Stage methods provided simpler yet consistent representations of trends. Information-theoretic comparisons using AIC confirmed that no single model consistently outperformed the others, and AIC-weighted multimodel averaging yielded stable and defensible quantile estimates that reduced dependence on individual models. Bootstrap resampling further highlighted the widening uncertainty bands for higher return periods, underscoring the challenges in estimating rare floods. However, the consistent upward signal across models suggests a robust intensification of flood hazard. Importantly, Stationary assumptions were found to overestimate present-day 50- and 100-year design floods by approximately 37–40%, which may lead to over-designed and cost-inefficient flood infrastructure if nonstationarity is ignored , implying significant risks of underdesign in hydraulic infrastructure. These findings underscore the importance of incorporating nonstationarity into FFA to support climate-resilient water resource planning, flood protection, and policy development. The methodological framework presented here can be extended by incorporating climate indices and land-use covariates into parameterizations, exploring Bayesian hierarchical models for probabilistic inference, applying the approach across multiple basins to capture regional patterns of risk, and integrating results into flood hazard mapping and early-warning systems.

Future research may extend the present framework in several directions. First, the incorporation of hydroclimatic and land-use covariates (e.g., large-scale climate indices, rainfall extremes, or catchment change indicators) into nonstationary model parameterizations could provide a more physically informed representation of flood drivers. Second, Bayesian hierarchical or fully Bayesian nonstationary models may be explored to better characterize parameter and structural uncertainty, particularly under

data-scarce conditions. Third, application of the proposed multimodel framework across multiple river basins would allow assessment of regional patterns in nonstationary flood behavior and improve the generalizability of the findings. Finally, integrating nonstationary design flood estimates into flood hazard mapping and risk-based infrastructure planning frameworks would enhance the practical utility of the approach for climate-resilient water resource management.

References

Barbhuiya, S., Ramadas, M. and Biswal, S.S. 2023. Nonstationary flood frequency analysis: review of methods and models. River, sediment and hydrological extremes: causes, impacts and management, 271-288.

Bhere, S. and Janga Reddy, M. 2025. Multi-covariate non-stationary flood frequency analysis for assessing flood hazard over the Mahanadi River basin, India. Hydrological Sciences Journal 70(1), 93-109.

Cheng, L., AghaKouchak, A., Gilleland, E. and Katz, R.W. 2014. Non-stationary extreme value analysis in a changing climate. Climatic change 127(2), 353-369.

Chow, K.A. and Watt, W.E. 1992. Use of Akaike information criterion for selection of flood frequency distribution. Canadian Journal of Civil Engineering 19(4), 616-626.

Coles, S., Bawa, J., Trenner, L. and Dorazio, P. (2001) An introduction to statistical modeling of extreme values, Springer.

Cui, H., Jiang, S., Gao, B., Ren, L., Xiao, W., Wang, M., Ren, M. and Xu, C.-Y. 2023. On method of regional non-stationary flood frequency analysis under the influence of large reservoir group and climate change. Journal of Hydrology 618, 129255.

de Bruijn, J.A., de Moel, H., Weerts, A.H., de Ruiter, M.C., Basar, E., Eilander, D. and Aerts, J.C. 2020. Improving the classification of flood tweets with contextual hydrological information in a multimodal neural network. Computers & Geosciences 140, 104485.

Debele, S., Strupczewski, W. and Bogdanowicz, E. 2017. A comparison of three approaches to non-stationary flood frequency analysis. Acta Geophysica 65(4), 863-883.

Dewan, T.H. 2015. Societal impacts and vulnerability to floods in Bangladesh and Nepal. Weather and Climate Extremes 7, 36-42.

Di Baldassarre, G., Laio, F. and Montanari, A. 2009. Design flood estimation using model selection criteria. Physics and Chemistry of the Earth, Parts A/B/C 34(10-12), 606-611.

Faulkner, D., Warren, S., Spencer, P. and Sharkey, P. 2020. Can we still predict the future from the past? Implementing non-stationary flood frequency analysis in the UK. Journal of Flood Risk Management 13(1), e12582.

Gül, G.O., Aşıkoğlu, Ö.L., Gül, A., Gülçem Yaşoğlu, F. and Benzeden, E. 2014. Nonstationarity in flood time series. Journal of Hydrologic Engineering 19(7), 1349-1360.

Haddad, K. and Rahman, A. 2011. Selection of the best fit flood frequency distribution and parameter estimation procedure: a case study for Tasmania in Australia. Stochastic Environmental Research and Risk Assessment 25(3), 415-428.

Hoq, M.S., Raha, S.K. and Hossain, M.I. 2021. Livelihood vulnerability to flood hazard: understanding from the flood-prone haor ecosystem of Bangladesh. Environmental management 67(3), 532-552.

Hounkpè, J., Diekkrüger, B., Badou, D.F. and Afouda, A.A. 2015. Non-stationary flood frequency analysis in the Ouémé River Basin, Benin Republic. Hydrology 2(4), 210-229.

Katz, R.W., Parlange, M.B. and Naveau, P. 2002. Statistics of extremes in hydrology. Advances in water resources 25(8-12), 1287-1304.

Khayyam, U. 2020. Floods: Impacts on livelihood, economic status and poverty in the north-west region of Pakistan. Natural Hazards 102(3), 1033-1056.

Kwon, H.-H., Sivakumar, B., Moon, Y.-I. and Kim, B.-S. 2011. Assessment of change in design flood frequency under climate change using a multivariate downscaling model and a precipitation-runoff model. Stochastic Environmental Research and Risk Assessment 25(4), 567-581.

Kwon, H.H., Brown, C. and Lall, U. 2008. Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modeling. Geophysical Research Letters 35(5).

López, J. and Francés, F. 2013. Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates. Hydrology and Earth System Sciences 17(8), 3189-3203.

Machado, M.J., Botero, B.A., López, J., Francés, F., Díez-Herrero, A. and Benito, G. 2015. Flood frequency analysis of historical flood data under stationary and non-stationary modelling. Hydrology and Earth System Sciences 19(6), 2561-2576.

Montello, F., Arnaudo, E. and Rossi, C. 2022. Mmflood: A multimodal dataset for flood delineation from satellite imagery. IEEE Access 10, 96774-96787.

Okoli, K., Breinl, K., Brandimarte, L., Botto, A., Volpi, E. and Di Baldassarre, G. 2018. Model averaging versus model selection: estimating design floods with uncertain river flow data. Hydrological Sciences Journal 63(13-14), 1913-1926.

Rajeshkannan, C. and Kogilavani, S. (2021) Mobile Computing and Sustainable Informatics: Proceedings of ICMCSI 2021, pp. 489-511, Springer.

Razmi, A., Golian, S. and Zahmatkesh, Z. 2017. Non-stationary frequency analysis of extreme water level: application of annual maximum series and peak-over threshold approaches. Water resources management 31(7), 2065-2083.

Strupczewski, W., Singh, V. and Feluch, W. 2001. Non-stationary approach to at-site flood frequency modelling I. Maximum likelihood estimation. Journal of Hydrology 248(1-4), 123-142.

Talchabhadel, R., Maskey, S., Gouli, M.R., Dahal, K., Thapa, A., Sharma, S., Dixit, A.M. and Kumar, S. 2023. Multimodal multiscale characterization of cascading hazard on mountain terrain. Geomatics, Natural Hazards and Risk 14(1), 2162443.

Zhang, T., Wang, Y., Wang, B., Tan, S. and Feng, P. 2018. Nonstationary flood frequency analysis using univariate and bivariate time-varying models based on GAMLSS. Water 10(7), 819.

Zhang, Z., Ma, Y. and Liu, P. 2024. A global multimodal flood event dataset with heterogeneous text and multi-source remote sensing images. Big Earth Data, 1-27.

Zheng, X., Duan, C., Chen, Y., Li, R. and Wu, Z. 2023. Disaster loss calculation method of urban flood bimodal data fusion based on remote sensing and text. Journal of Hydrology: Regional Studies 47, 101410.