Stochastic Environmental Research and Risk Assessment (2026) 40:124
https://doi.org/10.1007/s00477-026-03258-3

REVIEW PAPER

Intercomparison of regional flood frequency estimation procedures in West Africa

Serigne Bassirou Diop[1,2] · Yves Tramblay[2] · Ansoumana Bodian[1] · Bastien Dieppois[3] · Taha B. M. J. Ouarda[4]

Received: 9 March 2026 / Accepted: 23 April 2026

Abstract

West Africa faces devastating flood hazards that affect more than 400 million people. Yet flood risk assessment is hindered by sparse and often unreliable hydrological data. Regional flood frequency analysis (RFFA) is widely used to estimate design values at ungauged catchments and there is a need for a systematic intercomparison of RFFA approaches in this region. With an unprecedented dataset of 211 near-natural catchments, we compared a Direct Regression Approach (DRA) and three homogeneous region delineation methods using the index-flood methods based on spatial proximity, Principal Component Analysis (PCA), and Canonical Correlation Analysis (CCA) with catchment attributes. Each regional approach was paired with two regression models: (i) Stepwise Regression and (ii) Least Absolute Shrinkage and Selection Operator (LASSO), and four machine learning algorithms: (i) Random Forest (RF), (ii) eXtreme Gradient Boosting (XGB), (iii) Support Vector Regression (SVR), and (iv) a hybrid linear-tree ensemble (LinRF). Results show that index-flood methods consistently outperformed DRA, with the CCA-based framework achieving the highest accuracy. CCA-SVR combination is the best-performing regional model, yielding the lowest estimation errors (mean absolute relative error = 0.21 and relative bias = −0.03) for 20- or 50-year flood quantiles. Feature importance analysis revealed that subsurface properties, catchment area, and topographic attributes have stronger influence on regional flood estimation than surface features or land use patterns. The methodology and findings of this study offer practical tools for infrastructure design and climate adaptation, supporting more resilient flood risk management across vulnerable West African communities.

Keywords West Africa · Floods · Index-flood · GEV · CCA · Regionalization

1 Introduction

Floods are among the most devastating natural hazards worldwide (Delforge et al. 2025). Over the past few decades, the West African region has experienced an increase in flood frequency and severity (Ekolu et al. 2022; Nka et al. 2015; Tramblay et al. 2020; Wilcox et al. 2018), driven by a combination of climatic and non-climatic factors such as rapid and poor urban planning (Baddianaah 2023), land use changes (Hounkpè et al. 2019), extreme rainfall and soil moisture (Tramblay et al. 2021a), and limited adaptive capacity to climate change (Lalou et al. 2019; Niang et al. 2014; Roudier et al. 2011). Flood-related disasters have caused substantial loss of life, displacement of communities, and severe economic damage in this region (OCHA, 2025). This is particularly concerning given the region's ongoing rise in extreme rainfall events (Taylor et al. 2017; Tramblay et al. 2020), a trend projected to continue under future climate change storylines (Dosio et al. 2021; Chagnaud et al. 2023). Since extreme rainfall variability influences hydrological dynamics in the region (Elagib et al. 2021; Panthou et al. 2013), several studies have also reported increasing trends in extreme streamflow across West African catchments in both historical observations and future climate change projections (Aich et al. 2016; Diop et al. 2025; Ekolu et al. 2025; Nka et al. 2015; Wilcox et al. 2018). Effective

✉ Yves Tramblay
yves.tramblay@ird.fr

1 Laboratoire Leïdi 'Dynamique des territoires et développement', Université Gaston Berger (UGB), Saint Louis, Senegal

2 Espace-Dev, IRD, Univ. Montpellier, Montpellier, France

3 Centre for Agroecology, Water and Resilience (CAWR), Coventry University, Coventry, UK

4 Centre Eau Terre Environnement. INRS-ETE, 490 De la Couronne, Québec City, QC, Canada

Published online: 19 May 2026

flood risk assessment and management are therefore crucial to mitigate the socioeconomic and environmental impacts of flood hazards, and strengthen the regional resilience of West African communities.

Flood Frequency Analysis (FFA), a widely used engineering approach based on extreme value theory (Fisher and Tippett 1928), establishes the statistical relationship between flood magnitudes and their exceedance probabilities at a specific gauged site (Dalrymple 1960). FFA allows for design flood estimation, which is essential for planning and constructing hydraulic infrastructure (Feaster et al. 2023), and to support agricultural, industrial, and municipal water demands. However, the accuracy of these estimates depends on several interacting factors, including data sample size, extrapolation range, frequency model selection, calibration method, and non-stationarity, all of which can contribute to increased uncertainty in design flood estimation (Shimizu et al. 2020). Despite the urgency, flood risk assessment in West Africa remains hindered by persistent knowledge gaps, stemming from scarce and unreliable hydrological data. This limitation is especially critical when estimating rare events, as uncertainty increases with higher return periods (Wilcox et al. 2018), which require extrapolating beyond the available observations (Kidson and Richards 2005). In addition, small sample sizes may be insufficient for reliably estimating parameters of complex frequency models (Baidya et al. 2020). From a statistical point of view, a minimum of 30 data points is required to perform FFA (Douglas et al. 2000). Yet, the United States Geological Survey (USGS) recommends a minimum of 10 years of streamflow records to perform FFA (Bulletin 17 C guidelines; England Jr et al. 2019). Two main strategies are commonly applied to reduce the uncertainty associated with rare flood event estimation: (i) pooling flood data from multiple gauging stations within a defined region (Dalrymple 1960; Ouarda et al. 2008a, b); and (ii) incorporating extended records through historical (Hosking and Wallis 1986a) and paleohydrological data (Benito and O'Connor 2013; Hosking and Wallis 1986b). The first strategy, known as regional flood frequency analysis (RFFA) or regionalization, provides a robust framework for estimating flood quantiles at ungauged sites, addressing one of the most persistent challenges in catchment hydrology (Salinas et al. 2013). By pooling information from multiple sites, RFFA enhances the reliability of flood estimates, particularly in data-sparse regions.

RFFA approaches can be broadly classified into three main categories: (i) direct regression approaches (DRA) that relate catchment characteristics to flood quantiles; (ii) traditional index-flood method (Dalrymple 1960); and (iii) geostatistical techniques that leverage the spatial correlation of flood behavior (Chokmani and Ouarda 2004), often employing spatial interpolation or kriging methods. The dimensionless index-flood approach, first introduced by Dalrymple (1960), remains one of the most widely used methods for regional flood estimation. The index-flood method assumes that extreme floods across a homogeneous region share a common probability distribution, differing only by a site-specific scaling factor called the index-flood. It typically involves three steps: defining homogeneous regions, conducting at-site flood frequency analysis, and transferring regional information (Ouarda et al. 2001).

The increasing flood risk in West Africa contrasts sharply with the standards hitherto guiding hydrological design in the region. West African hydrological systems have undergone profound transformations over the past decades due to widespread land use changes and increasing climate variability (Mahé 2006; Sidibe et al. 2019; Herrmann et al. 2020). Yet, the two most used design flow guidance methods, the ORSTOM (Office de la Recherche Scientifique et Technique Outre-Mer in French, now the Institut de Recherche pour le Développement – IRD; Rodier and Auvray 1965) and CIEH (Comité Interafricain d'Études Hydrauliques in French; Puech and Chabi-Gonni 1983) guidelines, are based on hydrological data collected before 1965 and 1985, respectively. These legacy guidelines are no longer suited to the current hydrological dynamics in the region. Not only do they fail to account for the region's evolving environmental conditions, but they are also limited to 10-year flood estimations, making extrapolation to other return periods challenging and highly uncertain (FAO 1996). In addition, field evaluation revealed systematic biases of these two methods, when applied to recent observations in some West African watersheds (Agoungbome et al. 2018). Efficient water resources management depends on reliable estimates of flood frequency and magnitude to guide the design of hydraulic infrastructures (Feaster et al. 2023), as well as accurate runoff quantification to design optimal reservoirs for agricultural, industrial, and municipal water use (Song et al. 2022). Thus, inaccurate flood frequency estimations can lead to severe consequences, including dam failures, with far-reaching socioeconomic and environmental repercussions (Angnuureng et al. 2025; Beijing News, 2018; Kim and Lee 2020; Muhamad Bashar et al. 2025). Updating design value estimation guidelines is therefore imperative to ensure climate-resilient water management in the region (Amani and Paturel 2017).

Ahmed et al. (2023) conducted a worldwide bibliometric analysis and found a marked increase in the use of RFFA since 2007. However, their results were strongly biased by the expanding body of literature in Canada, the USA, the UK, Italy, and Australia. Very few applications of RFFA can be found for Africa. In North Africa, (Tramblay et al. 2024) made the first attempt to apply RFFA using the direct regression approach, and Morabbi et al. (2022) proposed a hydrological regions delineation method based on shifts in flood regime characteristics. In Southern Africa, Haile (2011) used the geographical grouping approach to perform RFFA based on the index-flood framework. Yet, West Africa represents a complex environment for implementing RFFA due to its climatic diversity (Nicholson 2018), and its pervasive data limitations (Tramblay et al. 2021b). In West Africa, the density of the hydrological monitoring networks have decreased in recent years (Tarpanelli et al. 2023). As a consequence, many stream-gauging sites have limited and often old flow data (Tramblay et al. 2021b), and numerous catchments remain ungauged. Furthermore, the presence of multiple transboundary river basins (World Bank 2021). and significant disparities in technical and institutional capacities across countries obstruct coordinated data management and hinder the development of harmonized regional flood models. As a result, to the best of our knowledge, RFFA studies in the West African context are limited only to the national or basin scale (e.g., Komi et al. 2016). These challenges highlight the value of RFFA as a tool for leveraging limited data across similar hydrological contexts, offering a pathway to support flood risk assessment and water resource planning across the region.

This is the first RFFA study focused on the whole West African region. The objective of the study is to test the feasibility of flood quantiles estimations at ungauged catchments, to inform infrastructure design and risk planning in a data-limited region. Appraisal studies of RFFA methodologies are often carried out based on rich databases with a large number of stations and long time series (see for instance: GREHYS 1996a, b). Very little effort has been dedicated to the evaluation of the robustness of these regional estimation procedures when applied under less-than-ideal conditions: limited number of gauged sites, low spatial density of stations, short time series, incomplete list of covariates, etc. The present study tries to address this limitation in a region where RFFA is intensely needed in hydrological modeling. The results of this study can eventually be of use in other regions of the world that are characterised by a notable lack of hydrological information.

First, we perform at-site flood frequency analysis by exploiting a recently compiled and unprecedented database, and second, we compare various regionalization approaches and regional transfer models. The present paper is organized as follows: after a description of the data bank in Sect. 2, the materials and methods used for delineation of homogeneous regions, and for regional estimation of flood quantiles are presented in Sect. 3. The findings are presented and discussed in Sect. 4, and finally, conclusions and perspectives are given in Sect. 5.

Table 1 Description and references of the catchment attributes included in the ADHI database

Description	Abbr.	References
Curve Number	CN	Ross et al. (2018) https://doi.org/10.1038/sdata.2018.91
Available Water Capacity	AWC	Wieder (2014) https://doi.org/10.3334/ORNLDAAC/1247
Bulk density	BD	
% Clay	Clay	
% Gravel	Gravel	
% Sand	Sand	
% Silt	Silt	
Groundwater depth	GD	MacDonald et al. (2012) https://doi.org/10.1088/1748-9326/7/2/024009
Groundwater productivity	GP	
Groundwater storage	GS	
Maximum altitude	MA	Lehner and Grill (2013) https://doi.org/10.1002/hyp.9740
Mean slope	Slope	
Mean altitude	Altitude	
Topographic Wetness Index	TWI	Sørensen et al. (2006) https://doi.org/10.5194/hess-10-101-2006
% Forest	Forest	ESA CCI LandCover http://www.esa-landcover-cci.org/
% Urban	Urban	
% Cropland	Cropland	
% Cropland irrigated	CI	
% Grassland	Grassland	
% Shrubland	Shrubland	
% Sparse	Sparse	
% Bare land	BL	
Area	Area	
Mean annual precipitation	Prec	Harris et al. (2020) https://doi.org/10.1038/s41597-020-0453-3
Mean annual temperature	Temp	
Mean annual potential evapotranspiration	PET	

2 Hydrological database

We use the hydrological observation data from the African Database of Hydrometric Indices (ADHI; Tramblay et al. 2021b). This comprehensive database compiles key hydrological indicators derived from various sources, with daily discharge records extending over a minimum of 10 years between 1950 and 2018. In addition to discharge data, the database includes a range of physiographic characteristics describing the associated watersheds. Table 1 lists all variables available in the ADHI database with their references. The catchment descriptors include standard physiographical properties that are routinely used in RFFA. The selection of specific flood predictor variables for the analysis (Sect. 3.2.1) is designed to provide a broad characterization of basin properties rather than a priori prioritization, ensuring that the dataset encompasses most physiographical factors that could influence flood response. This approach follows the practice in previous RFFA studies, where comprehensive variable sets are used to capture inter-basin variability (GREHYS 1996a, b; Ouarda et al. 2001; Dawson et al. 2006; Preti et al. 2011; Haddad and Rahman 2012; Rahman et al. 2018; Han et al. 2020; Msilini et al. 2022; Tramblay et al. 2024; Fitriana et al. 2025).

Only unregulated catchments, determined through visual inspection of dam locations relative to watershed outlets, with at least 10 years of daily data for the period 1975–2018, and a drainage area smaller than 150,000 km² were selected. Based on these criteria, a total of 211 streamflow gauges were retained for the analysis. Figure 1 shows the geographical distribution of the selected sites. We applied the block-maximum (Gumbel 1958) approach, to extract annual maximum flow (AMF) values from daily discharge time series. To avoid missing true flood peaks, years with data gaps near the peak flow were excluded after visually inspecting each station's yearly hydrograph. It is worth noting that the AMF series of selected sites are independent and stationary, making them well-suited for FFA.

3 Materials and methods

3.1 At-site flood frequency analysis

Local flood frequency analysis is generally performed in two steps.First, the most appropriate regional probability distribution is identified using goodness-of-fit tests, such as the Anderson-Darling and Kolmogorov-Smirnov statistical tests, and the Akaike Information Criterion and the Bayesian Information Criterion. Then, flood quantiles are estimated individually for each site across the study area. Since there is no universally accepted model for flood frequency analysis (GREHYS 1996a), evaluating multiple candidate probability distributions and parameter estimation methods is essential to identify the best-suited frequency model for a given dataset. Diop et al. (2025), considering both stationary and non-stationary modelling frameworks, revealed that the Generalized Extreme Value (GEV) with the Generalized (Penalized) Maximum Likelihood (GML; Martins and Stedinger 2000) parameter estimation method provided the best results. Therefore, here, we use the GEV-GML model for estimating flood quantiles. The cumulative distribution function (CDF) of the GEV model is presented in Eq. (1). The GEV model is a continuous distribution with three parameters (i.e., location, scale, and shape). From Eq. (1), the shape parameter (ξ) defines the tail behavior of the GEV distribution, which unifies three distinct types of extreme value distributions (Coles 2001): (i) when $\xi > 0$, the distribution corresponds to the heavy-tailed Fréchet type (Fréchet 1927), (ii) when $\xi = 0$, the GEV reduces to the light-tailed Gumbel distribution (Gumbel 1958); and (iii) when $\xi < 0$, it represents the bounded Weibull type (Weibull 1951), indicating an upper limit on flood magnitudes.

$$F(x;u,\alpha,\xi) = exp\left\{-\left[1-\xi\,\frac{x-u}{\alpha}\right]^{1-\xi}\right\} \quad \xi \neq 0 \tag{1}$$

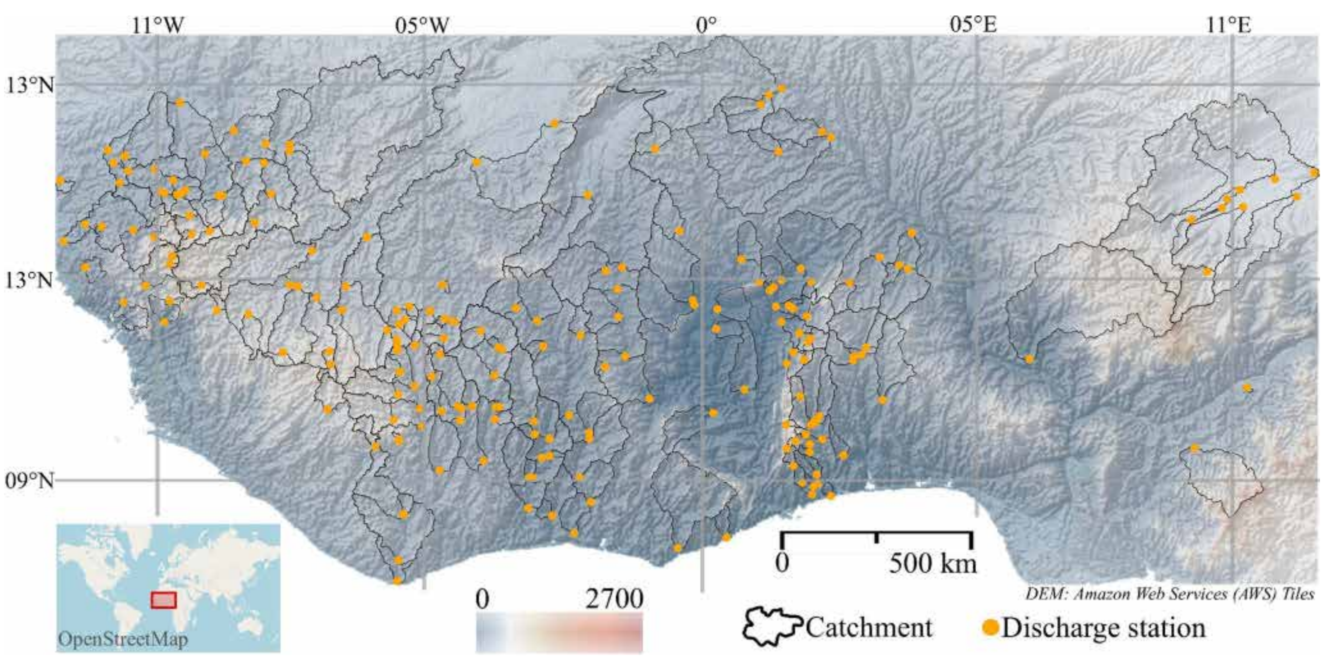

Fig. 1 Location of discharge gauging stations (orange dots) selected for the study, with the associated drainage basins (black contours), and discharge stations (orange points). The color shading represents topographic elevation, with darker blues indicating low-lying areas and reddish tones highlighting higher terrain

$$F(x;u,\alpha) = exp\left\{-exp\left[-\frac{x-u}{\alpha}\right]\right\} \qquad \xi = 0$$

Where x, u, α and ξ are the data, location, scale and shape parameters respectively, and $(u+\alpha/\xi) \le x < \infty$ if $\xi < 0$; $-\infty < x < \infty$ if $\xi = 0$; $-\infty < x \le (u+\alpha/\xi)$ if $\xi > 0$.

The Maximum Likelihood (ML) method is widely accepted as a robust method for the estimation of frequency models' parameters. Yet, the ML method can be unstable with small sample sizes, particularly when estimating the GEV shape parameter, often yielding unrealistic shape values. Therefore, here, we use the GML method to overcome this issue by incorporating a beta prior distribution (with shape parameters $p = 6$ and $q = 9$) to constrain the GEV shape parameter within the interval [–0.5, +0.5] (El Adlouni et al. 2007; Martins and Stedinger 2000). In addition, as in Diop et al. (2025), the beta prior was adapted for the West African context, using a normal distribution with a mean of –0.24 and a standard deviation of 0.16, better reflecting regional climatic characteristics.

3.2 Regional flood frequency analysis

3.2.1 Selecting flood predictor variables

The relationship between watershed characteristics, such as physiographic, meteorological, and geological features and flood generation processes is inherently complex. This complexity presents significant challenges when attempting to delineate hydrologically homogeneous regions based solely on catchment attributes (Ahani et al. 2018). Moreover, the development of robust regression models for regional flood estimation is often hampered by high heterogeneity among catchments and weak relationships between predictors and hydrological responses. Feature selection (FS) techniques offer a solution to address these challenges, by removing non-informative and redundant variables, thereby enhancing model interpretability and reducing overfitting without substantial loss of predictive information (Cai et al. 2018; Muthukrishnan and Rohini 2016). FS facilitates the identification of the most influential catchment properties driving flood behavior, which is critical for both building reliable regression models and delineating hydrologically homogeneous regions. FS methods are commonly categorized into filter and wrapper approaches (Kuhn and Johnson 2013). Filter methods operate independently of any learning algorithm, relying on statistical criteria (e.g., Spearman correlation and F-statistic) to rank and select variables. Filter methods are computationally efficient, less prone to overfitting, and model-agnostic, but they ignore interactions among features and may retain redundant predictors. Conversely, wrapper methods evaluate subsets of features based on model performance, allowing them to capture feature interactions more effectively. However, they are model-specific, computationally intensive, and more susceptible to overfitting, particularly in data-scarce contexts.

Here, we apply the minimum redundancy maximum relevance (MRMR) algorithm to the entire dataset as a preliminary step, before any further analysis. MRMR is an FS method that seeks to balance predictive relevance with low inter-feature redundancy (Ding and Peng 2005). The MRMR algorithm maintains the efficiency and model-independence of filter methods while mitigating redundancy. Although it may not always achieve the performance gains of wrapper methods tailored to specific models, the MRMR algorithm provides a robust and computationally efficient compromise, particularly suited for large feature spaces, and reduces overfitting risk. The MRMR algorithm iteratively constructs a set of k-optimal features. In the first iteration, it selects the variable with the highest relevance to the target (quantified here using the F-statistic). At each subsequent step, it chooses the next feature that maximizes relevance to the target while minimizing redundancy with the previously selected features. This balance is formalized through the MRMR objective function given in Eq. (2). The optimal number of features to retain is a tunable hyperparameter, determined in the present work via 10-fold cross-validation using a multiple linear regression as predictive model and the Root Mean Square Error (RMSE) as scoring function. This ensures that the selected subset not only maximizes the MRMR optimization criteria, but also performs well in a predictive modeling context.

$$f(x) = \frac{F(y,x_i)}{\frac{1}{S}\sum_{j=1}^{S}\rho(x_i,x_j)} \tag{2}$$

$$F(y,x_i) = \frac{\rho(y,x_i)^2}{\left[1-\rho(y,x_i)^2\right]}\cdot(n-2)$$

From Eq. (2), $f(x)$ is the MRMR score of feature x_i, $F(y,x_i)$ represents the F-statistic which represent the relevance score, $\rho(y,x_i)$ is the Pearson correlation coefficient between the target variable y and the feature x_i; $\rho(x_i,x_j)$ represents the Pearson correlation coefficient between feature x_i and each already selected feature x_j, S is the number of features already selected, and $n-2$ is the degrees of freedom for a simple linear regression with n samples, one predictor, and an intercept.

3.2.2 Delineation of homogeneous regions

Homogeneous regions for regional flood frequency analysis can be delineated as geographically contiguous or not (fixed-region method), or based on hydrological neighborhood (Ouarda 2016). Neighborhood approaches are particularly flexible and include methods such as the Region of Influence (ROI) approach (Burn 1990) and the Canonical Correlation Analysis (CCA) method (Ouarda et al. 2001). In the fixed-region method, homogeneous regions are predefined and often consist of a set of sites located within a static geographical boundary (Burn 1988). In contrast, the neighborhood (or site-focused) approach treats each site as the "center" of its own tailored region. The site-focused approach often yields more efficient and accurate flood quantile estimates than fixed-region methods (Haddad and Rahman 2012), particularly in heterogeneous landscapes where rigid regional boundaries may fail to adequately represent the true hydrological similarity of a target basin. The neighborhood method defines a unique neighborhood (or set of donor sites) for each target site based on proximity within a multidimensional attribute space, which may include geographical coordinates as well as climatic and physiographic attributes (GREHYS 1996b). Here, proximity is computed in both geographical space (from coordinates) and physiographical space via Principal Component Analysis (PCA) and CCA.

PCA is a dimensionality reduction technique that projects high-dimensional data onto a set of orthogonal components (linear combinations of the original variables) ranked by the variance they explain (Dorabiala et al. 2024; Younes et al. 2023). Like many variance-based algorithms, PCA is sensitive to feature scale. To avoid biasing the influence of variables with larger magnitudes (Gewers et al. 2022), we apply PCA to standardized data (i.e., features with zero mean and unit variance). This preprocessing preserves the data structure and allows an unbiased comparison of feature contributions. CCA is a statistical multivariate technique that identifies pairs of orthogonal linear combinations (canonical variates) from two variable sets to maximize their mutual correlation (Muirhead 1982). CCA allows the identification of hydrologically similar gauged sites for an ungauged target basin, by relating catchment attributes to flood quantiles. CCA has been widely applied in regional flood frequency analysis (Cavadias 1989, 1990; Chebana and Ouarda 2021; Jung et al. 2019; Ouarda et al. 2001; Ribeiro-Corréa et al. 1995; Shu and Ouarda 2007).

A critical aspect of the neighborhood approach is balancing the size of the pooling group with its internal homogeneity. While larger groups enhance reliability in flood quantile estimation, they may introduce heterogeneity. Therefore, it is essential to develop objective criteria for determining the optimal number of donors to address this trade-off (Ouarda et al. 2001). We therefore adopt the 5T rule from the Flood Estimation Handbook (Robson and Reed 1999), which requires that pooled sites collectively provide at least five times the target return period in station-years. For both geographical and PCA spaces, Euclidean distances between the target and potential donor sites are computed and sorted to prioritize similarity. Neighbors are then identified iteratively: sites are added one by one until the 5T threshold is met. For the CCA approach, Ouarda et al. (2001) provided a detailed algorithm to determine the optimal number of donor sites for a given target location. However, this method requires the optimization of a parameter that defines the region size, which can be challenging, especially in data-scarce contexts. Thus, here, we use the modified version proposed by (Girard 2001), which is free of parameter tuning, making it less subjective and easier to apply (Ouarda et al. 2008a, b).

3.2.3 Homogeneity assessment

We have applied two statistical tests to assess the homogeneity of the delineated regions. To avoid data leakage, the target site data is omitted from the regional test. For the geographical and PCA-based proximity, the homogeneity testing is performed iteratively using the Hosking and Wallis (HW) heterogeneity test (H-statistic; Hosking and Wallis 1993). At each step, homogeneity is assessed, and sites that compromise regional homogeneity are excluded. The H-statistic compares the variability of site L-moment ratios against expectations from a homogeneous region, simulated herein via 1000 Monte Carlo samples drawn from a fitted Kappa distribution. The Kappa distribution, whose cumulative distribution function (CDF) is given in Eq. (3), is a four-parameter continuous distribution that can simplify to several commonly used distributions, including the Generalized Logistic (GLO), Generalized Pareto (GPA), and GEV distributions. For the CCA method, the HW test was first performed for each given region; if the test fails, we then use the Hosking and Wallis discordance measure (D-measure; Hosking and Wallis 1993) to remove problematic sites. The D-measure identifies sites that differ significantly from the pooled group based on their L-moment ratios. Excluding such sites improves regional homogeneity (Hosking and Wallis 1993).

$$F(x;\mu;\alpha;k;h) = \left\{1 - h\left[1 - k\frac{x-\mu}{\alpha}\right]^{1/k}\right\}^{1/h} \tag{3}$$

From Eq. (3), x, u, α, k and h denote the data, location, scale, and the first and second shape parameters, respectively. The

Kappa distribution reduces to several well-known distributions for specific values of h and k: when h = –1 it becomes a GLO; when h=0, it reduces to the GEV; when h=1, it simplifies to the GPA; and for h=1 and k=0, it corresponds to the Exponential distribution.

3.2.4 Regional estimate of flood quantiles

Figure 2 illustrates the computational workflow adopted in this study for regional flood quantile estimation. To assess whether a global model is sufficient or whether regionalization based on hydrologically homogeneous regions is required, we evaluated a Direct Regression Approach (DRA) as a benchmark against three delineation methods implemented within the index-flood framework: spatial proximity, PCA, and CCA. The index-flood approach involves delineating hydrologically homogeneous regions before regional model development, in which the 211 basins are partitioned according to site similarity criteria derived from spatial, physiographic, and hydroclimatic characteristics. Both the DRA and the index-flood-based regionalization approaches were applied using two regression models: (i) Stepwise regression (STEP) and (ii) Least Absolute Shrinkage and Selection Operator (LASSO), as well as four machine learning (ML) algorithms: (iii) Random Forest (RF), (iv) eXtreme Gradient Boosting (XGB), (v) Support Vector Regression (SVR), and (vi) a hybrid linear-tree ensemble (LinRF). Within the index-flood framework, these regression/ML models were used to estimate the index flood (i.e., the mean annual flood) from catchment descriptors, while a regional growth curve was derived independently for each delineated region. Conversely, under the DRA, the same models were applied to estimate flood quantiles directly from catchment descriptors, thereby bypassing the index-flood scaling procedure.

3.2.4.1 The index-flood framework Estimating flood quantiles at ungauged sites using the index-flood method involves scaling data from gauged sites within a region by a site-specific factor before any regional model building. The site-specific factor, or index flood, is typically represented by the mean or median of the annual maximum flow series at each site when data are available. Here, we use the sample mean, which is commonly used in the literature as the scale factor (Chebana and Ouarda 2009; Dalrymple 1960; Hosking and Wallis 1993; NERC 1975). For completely ungauged catchments, the index flood is predicted using regional regression or machine learning models that relate hydrological responses to catchment characteristics from gauged sites (Pandey and Nguyen 1999). To prevent

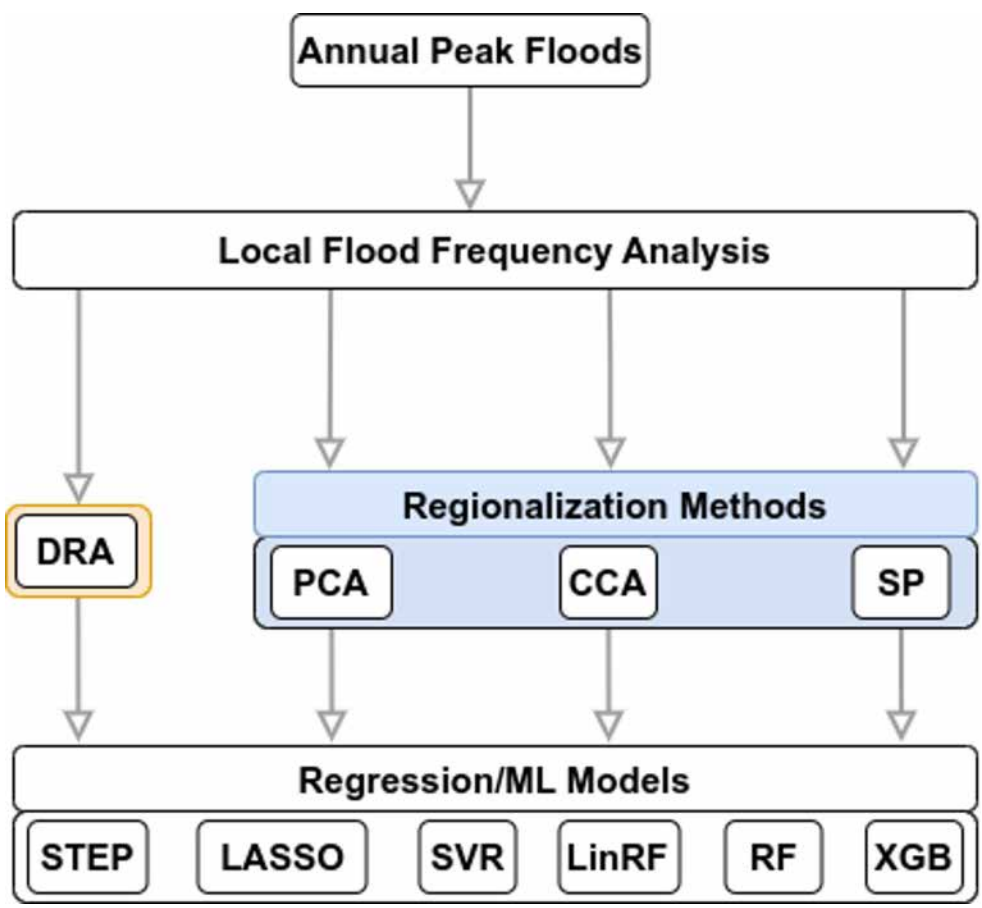

Fig. 2 Workflow adopted in the present RFFA. The process integrates local flood frequency analysis using GEV and Gumbel distributions, with model selection based on goodness-of-fit criteria (AIC/BIC), together with global (DRA) and regional (Index-Flood) approaches. Homogeneous regions for the index-flood framework are identified using PCA, CCA, or spatial proximity (SP) methods. A suite of regression/ML models including—(i) Stepwise regression (STEP), (ii) LASSO, (iii) Random Forest (RF), (iv) eXtreme Gradient Boosting (XGB), (v) Support Vector Regression (SVR), and (vi) a hybrid linear-tree Random Forest (LinRF)—is used to relate catchment attributes to flood quantiles directly for the DRA, and to scaling factors for the index-flood-based approaches

bias and data leakage, we implement a jackknife resampling approach: for each target site, its flood data are excluded, and the model is trained using the remaining gauged sites. The calibrated model then predicts the index flood for the target site based on its physical and climatic attributes. It is important to note that the index flood (mean annual flood) is always estimated locally for each site. Regional regression or machine learning models are only used to predict the index flood at ungauged sites. Once the site-specific index floods are obtained, the annual maximum flood series from all gauged sites are scaled by their respective index floods to produce a pooled, dimensionless series. This pooled series is then used to derive a regional growth curve. Finally, the flood quantile at a given return period for an ungauged site is estimated using the classical index-flood relationship (Eq. 4). This workflow ensures that only the growth curve is regionalized, while the mean annual flood remains a site-specific estimate.

$$Qreg = Gt \cdot Qi \tag{4}$$

From Eq. (4), Qreg is the regional flood quantile, Gt is the growth factor for an annual exceedance probability of 1 in t, and is the site-specific index flood (mean annual peak flow).

3.2.4.2 Regression and ML algorithms Linear regression models are prone to overfitting when the number of predictors is large, especially in the presence of multicollinearity. To address the high dimensionality of basin descriptors, we have evaluated the STEP and LASSO multiple linear regression models. STEP and LASSO regressions are commonly used for feature selection, relying on the statistical significance (e.g., *p*-values) of predictors. These techniques reduce overfitting by retaining only the most informative features, thereby improving generalization. Stepwise regression incrementally adds or removes variables based on their contribution to model performance. There are three types of stepwise regressions: (i) forward selection, which begins with no predictors and sequentially adds the most statistically significant ones; (ii) backward elimination, which starts with all predictors and iteratively removes the least significant ones based on their impact on model performance; and (iii) bidirectional elimination, which combines forward selection and backward elimination in a hybrid approach. In this study, we apply the backwards elimination strategy. LASSO regression (Tibshirani 1996) is a regularization-based technique that simultaneously performs feature selection and prevents overfitting by shrinking some coefficients towards zero. This is achieved by adding a penalty term (λ) to the regression loss function (Ranstam and Cook 2018). A higher penalty results in fewer features being selected, as it forces more coefficients to shrink to zero, while a lower penalty retains more features by reducing the penalty effect. The main advantage of LASSO over the stepwise model is that it can better handle multicollinearity.

SVR (Drucker et al. 1996) is an extension of the kernel-based Support Vector Machine (SVM) classification algorithm (Vapnik 1995) for regression tasks. Unlike traditional linear regression models that aim to minimize the overall prediction error, SVR introduces the concepts of hyperplanes and margins to define an ε-insensitive region (or ε-tube). Within this tube, prediction errors smaller than ε are not penalized. Only data points lying outside this region (known as support vectors) influence the optimization process and contribute to the loss function. The trade-off between model complexity and the tolerance for deviations within the ε-insensitive zone (ε-tube) is governed by a regularization parameter (Schölkopf et al. 1998), which helps mitigate overfitting. We use a linear SVR herein with a linear kernel, assuming a linear relationship between input predictors and the target variable.

A potential limitation of linear models in RFFA is their assumption of linearity, whereas hydrological processes often exhibit complex non-linear behaviors (Msilini et al. 2020; Sivakumar and Singh 2012). Tree-based models, such RF and XGB, have emerged as effective alternatives, offering greater flexibility in capturing interactions and non-linear relationships (Carvalho et al. 2018). Tree-based model predictions are not continuous but represent constant approximations of the target variable at each terminal node. As a result, they struggle with linear relationships (Desai and Ouarda 2021) and are highly sensitive to small changes in input data. This makes them prone to overfitting, especially when their growth is not properly regulated or when test samples fall into underexplored regions of the feature space (Zhang et al. 2019). To overcome these limitations, here, we also employ LinRF, a hybrid approach combining the RF algorithm with multiple linear regression. LinRF employs multiple linear regression models instead of constant values for the prediction of leaf nodes (Ao et al. 2019; Cerliani 2021). We also use the XGB algorithm with a gradient boosting decision tree (GBDT), which sequentially builds decision trees to minimize prediction errors (Chen and Guestrin 2016). In GBDT, each new learner addresses the weak points of the previous prediction model by approximating pseudo-residuals, refining the model's accuracy. GBDT also incorporates L1 (λ) and L2 (α) regularization to control complexity and reduce overfitting, making it a flexible predictive tool.

3.2.4.3 Hyperparameters optimization Table 2 presents the sets of hyperparameters used for the RF and XGB algorithms. Hyperparameter tuning was performed using randomized search within a cross-validation framework to identify optimal model configurations while limiting overfitting. For the RF algorithm, hyperparameter tuning explored ranges of 10–500 trees (n_estimators), 5–25 maximum leaf nodes (max_leaf_nodes), 5–25 minimum samples per leaf (min_samples_leaf), and 5–15 minimum samples required for node splitting (min_samples_split). For XGB, hyperparameter tuning explored 1–200 boosting rounds (n_estimators), and L1 regularization term (alpha) ranging from 10^{-4} to 10, and a log-uniform learning rate (learning_rate) between 0.01 and 0.5.

3.2.5 Evaluation criteria

We have performed a leave-one-out (jackknife) sampling, where each site is alternately treated as ungauged, and models are trained on the remaining data. The jackknife approach allows site-specific error evaluation but may introduce high variability due to reliance on single-point testing (Luntz and Brailovsky 1969). Therefore, we also have implemented K-Fold cross-validation (K-cv) within the DRA to provide a more robust assessment of model performance. In K-cv, the dataset is randomly split into K-equal subsets (k = 100 herein), with the model trained on all but one fold and tested on the remaining one (James et al. 2017). K-cv allows assessment of the model's performance across different data distributions. We compute two scale-independent metrics for both resampling strategies: (i) relative bias (rBias), which quantifies systematic over- or underestimation, and (ii) mean absolute relative error (MARE), which represents the average magnitude of prediction errors.

$$MARE = \frac{1}{n}\sum{}_{i=1}^{n}\left|\frac{P_i - O_i}{O_i}\right| \tag{5}$$

$$rBias = \frac{1}{n}\sum{}_{i=1}^{n}\frac{P_i - O_i}{O_i} \tag{6}$$

From Eqs. (5) and (6), $MARE$ and $rBias$ are respectively the mean absolute relative error and relative bias, n represents the sample size, P_i and O_i are respectively the predicted and the actual flood quantiles at the station i.

Table 2 Sets of optimal hyperparameters used for the RF and XGB algorithms

ML algorithm	Hyperparameter	Optimal value
XGB	n_estimators	100
	alpha	0.1
	learning_rate	0.29
RF	max_leaf_nodes	42
	min_samples_leaf	5
	min_samples_split	6
	n_estimators	500

4 Results and discussions

4.1 The relation between flood quantiles and catchment attributes

To identify relevant features for RFFA, we have applied the MRMR algorithm to the full dataset as an initial screening method. The optimal number of predictors is treated as a hyperparameter, optimized via 10-fold cross-validation. A multiple linear regression model was used, with the 20-year flood quantile as the target variable and watershed physiographic and climatic attributes as predictors. The 20-year flood is commonly used as a benchmark in flood frequency analysis (Franks and Kuczera 2002; Dawson et al. 2005; Hamlet and Lettenmaier 2007; Reynard et al. 2004; Arnell and Gosling 2016; Yang et al. 2016; Kay et al. 2018; Tramblay and Somot 2018; Smith and Alexander 2018; Hu et al. 2020; Han et al. 2022; Eslamian et al. 2025), providing an optimal balance between the scarcity of extreme events (limited by record length) and the statistical reliability of quantile estimates. Model performance is evaluated by the RMSE metric. Optimal predictive performance was obtained with a subset of 14 predictors, as shown in Fig. 3a. Catchment area (Area) ranks highest in MRMR scores, followed by the topographic wetness index (TWI), maximum altitude (MA), and geological descriptors such as groundwater depth (GD), groundwater productivity (GP), and groundwater storage (GS) (Fig. 3b). The MRMR results are consistent with the results from the Spearman's rank correlation analysis (Fig. S1), which examines the relationship between watershed attributes and two flood metrics: the 20-year flood quantile and the mean annual maximum flood (AMF). The strongest correlation was found with catchment area ($\rho = 0.77$), followed by geological properties GD and GP ($\rho = 0.73$ and $\rho = 0.68$, respectively), TWI ($\rho = 0.65$), and maximum altitude MA ($\rho = 0.62$). All of these correlations are statistically significant at the 0.05 level. This suggests that catchment attributes are potential drivers of floods in the study area. While the catchment area may influence both the volume and timing of flood events (Stein et al. 2021), the geological characteristics of a watershed modulate groundwater storage and significantly affect both surface runoff and baseflow (Rosbjerg et al. 2013).

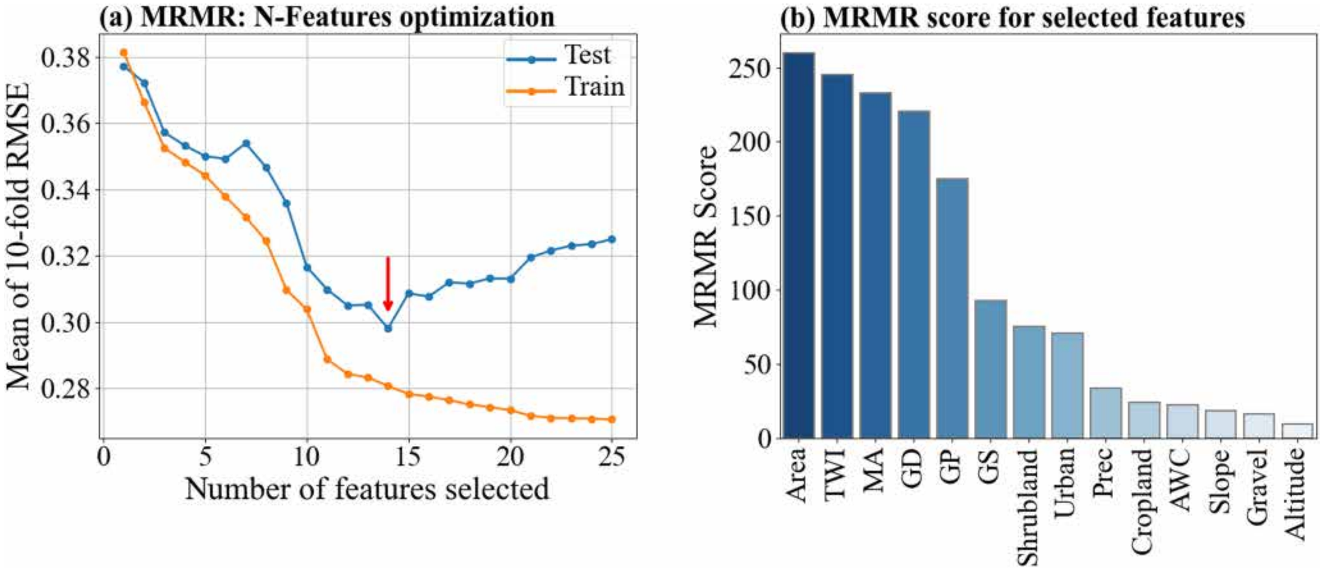

Fig. 3 Feature selection using the MRMR algorithm: **a** Optimization of the number of selected features based on the mean 10-fold cross-validation RMSE; **b** MRMR scores for the selected features. The orange and blue lines show the training and test RMSE, respectively, both as functions of the number of features retained

4.2 Delineation of hydrologically homogeneous regions

The PCA and CCA approaches are first carried out on the whole dataset to further assess the contribution of each catchment attribute before the identification of homogeneous regions for flood quantiles estimation. Figure 4 shows a summary of the PCA results with the projection of the selected catchment attribute onto the first two principal components (PC1 and PC2; Fig. 4a), and the relative contribution of each attribute to the variability captured by the individual principal components (Fig. 4b). PC1 is strongly associated with catchment area (Area), TWI, and geological attributes (GP, GD, and GS), indicating that these variables tend to vary together in the same direction. PC2 is primarily driven by Prec, MA, Altitude, and Gravel, suggesting it reflects gradients in climatic and elevation characteristics. PC3 captures variability mainly linked to slope and shrubland coverage, while PC4 is dominated by the AWC variable and proportion of urban areas. The first four principal components explain almost 80% of the variation in the dataset. CCA is conducted using the 14 selected catchment attributes as psychographic variables, while the 5-year flood and the Q50/Q5 ratio are used as the hydrological response variables. According to Ribeiro-Corréa et al. (1995), the interpretation of canonical variables is best achieved by examining their correlations with the catchment attributes. Table 3 gives the correlation between the two physiographic canonical variates and the catchment attributes. The most influential features contributing to the interpretation of the canonical variables include Area, GD, TWI, precipitation (Prec), GS, GP and MA, indicating a strong association with catchment size, geology properties, precipitation gradients and wetness conditions (Table 3). Regarding the delineation of homogeneous regions, the number of neighbors varies by method. For the spatial proximity approach, the number of neighboring stations ranges from 19 to 33, with an average of 26 sites per region. For the PCA-based method, regions include between 14 and 32 donor sites, with a mean of 18 neighbors, and for the CCA approach, the number of neighbors varies between 13 and 75, with a mean region size of 51 donors.

4.3 Regional estimation of flood quantiles

Here, we contrast the DRA, considered as a global benchmark model, to three regional delineation methods within the index-flood framework. This setup allows the assessment of both regionalized (index-flood) and global (DRA) approaches, resulting in a total of 24 model configurations.

4.3.1 Accuracy of regional flood quantile estimates compared to local estimates

Figure 5 shows a comparison between regional and local estimates for 20- and 50-year floods, based on a jackknife procedure. The accuracy of regional flood quantile estimates varies depending on the model configuration. The DRA shows the poorest performance overall, particularly with linear models (e.g., STEP, LASSO), which tend to systematically underestimate high flood magnitudes (Fig. 5a). In contrast, index-flood-based methods provide significantly

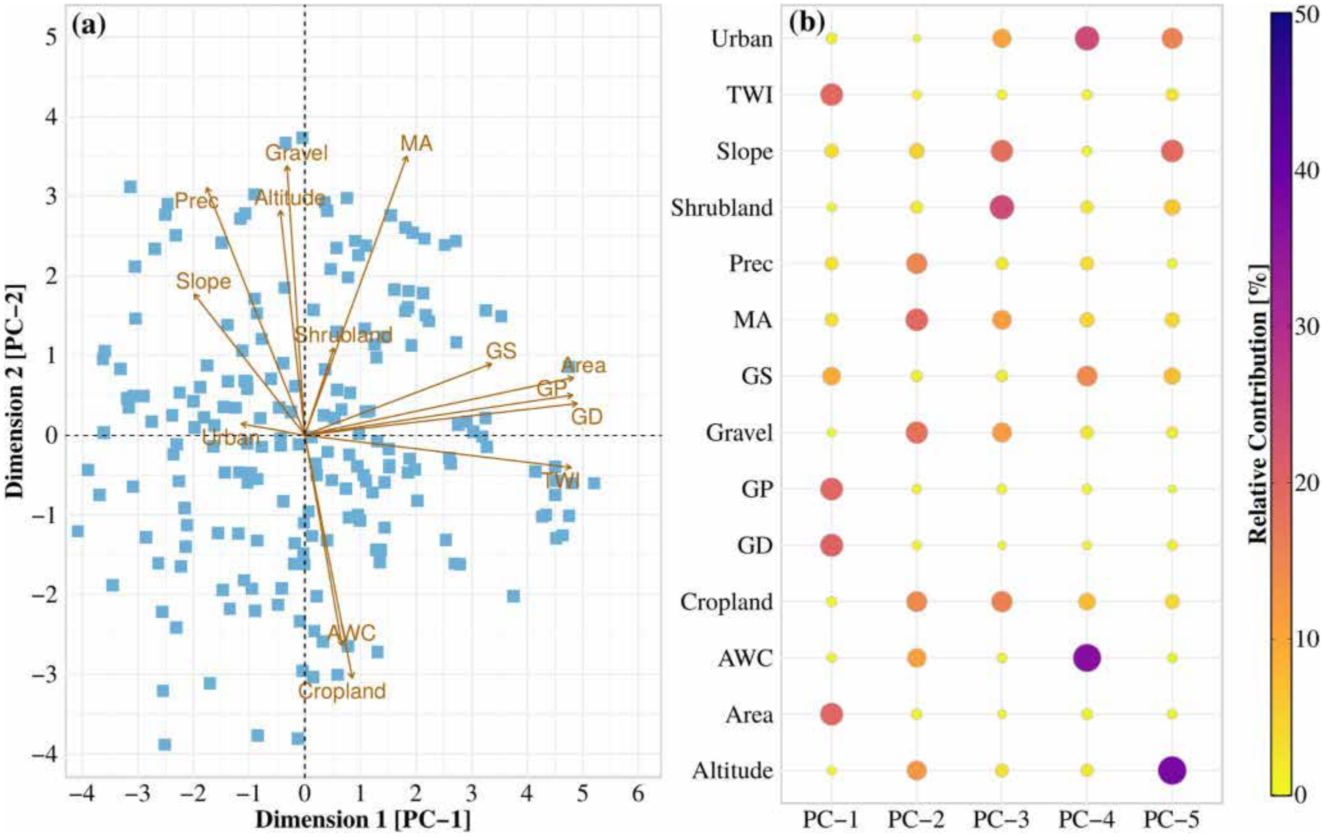

Fig. 4 PCA biplot (**a**) showing the projection of the selected catchment attributes onto the first two principal components (PC1 and PC2), and relative contribution of each variable to the first five principal components (**b**)

Table 3 Correlation between physiographic canonical variables and catchments attributes

Original variables	Can. variable 1	Can. variable 2
TWI	0.65	0.007
Area	0.62	0.16
GD	0.37	0.05
Prec	0.33	0.15
GS	0.30	0.03
GP	0.23	0.11
MA	0.23	0.17
AWC	0.14	0.21
Slope	0.12	0.39
Cropland	0.11	0.02
Gravel	0.10	0.10
Urban	0.10	0.12
Shrubland	0.08	0.10

improved performance, as indicated by regional estimates that more closely align with the corresponding local flood quantiles along the 1:1 reference line. Unlike the index-flood framework, which accounts for regional heterogeneity through pooling of hydrologically similar catchments, the DRA treats the entire study area as a single region and directly relates physio-climatic attributes to flood quantiles. This makes the DRA sensitive to the heterogeneity in catchment size, geology, and rainfall regimes, often leading to biased estimates (Salinas et al. 2013). Among neighborhood approaches, the CCA method, designed to maximize the correlation between catchment descriptors and flood quantiles, provides the least dispersed scatter plots, regardless of the regression or ML model. Regarding regression and ML models, tree-based algorithms (RF and XGB) show the widest spread of regional-to-local discrepancies on the scatter plots among all models. Interestingly, the LinRF consistently outperforms the classical RF, highlighting its effectiveness in learning from limited sample sizes (Ao et al. 2019).

It is worth noting that in Fig. 5, the DRA shows a slightly lower coefficient of determination (R^2) for the 20-year flood compared to the 50-year flood, even though the MARE tends to be higher for larger return periods (Fig. 6). This pattern may arise from the properties of R^2, which depends on the variance of the observed values. For larger return periods (e.g., 50-year), inter-basin variability of flood magnitudes is substantially higher, increasing the total variance in the denominator of R^2. Consequently, even if absolute residuals are larger, the proportion of variance explained by the model can also increase, yielding higher R^2 values. In contrast, at lower return periods (e.g., 20-year), the lower

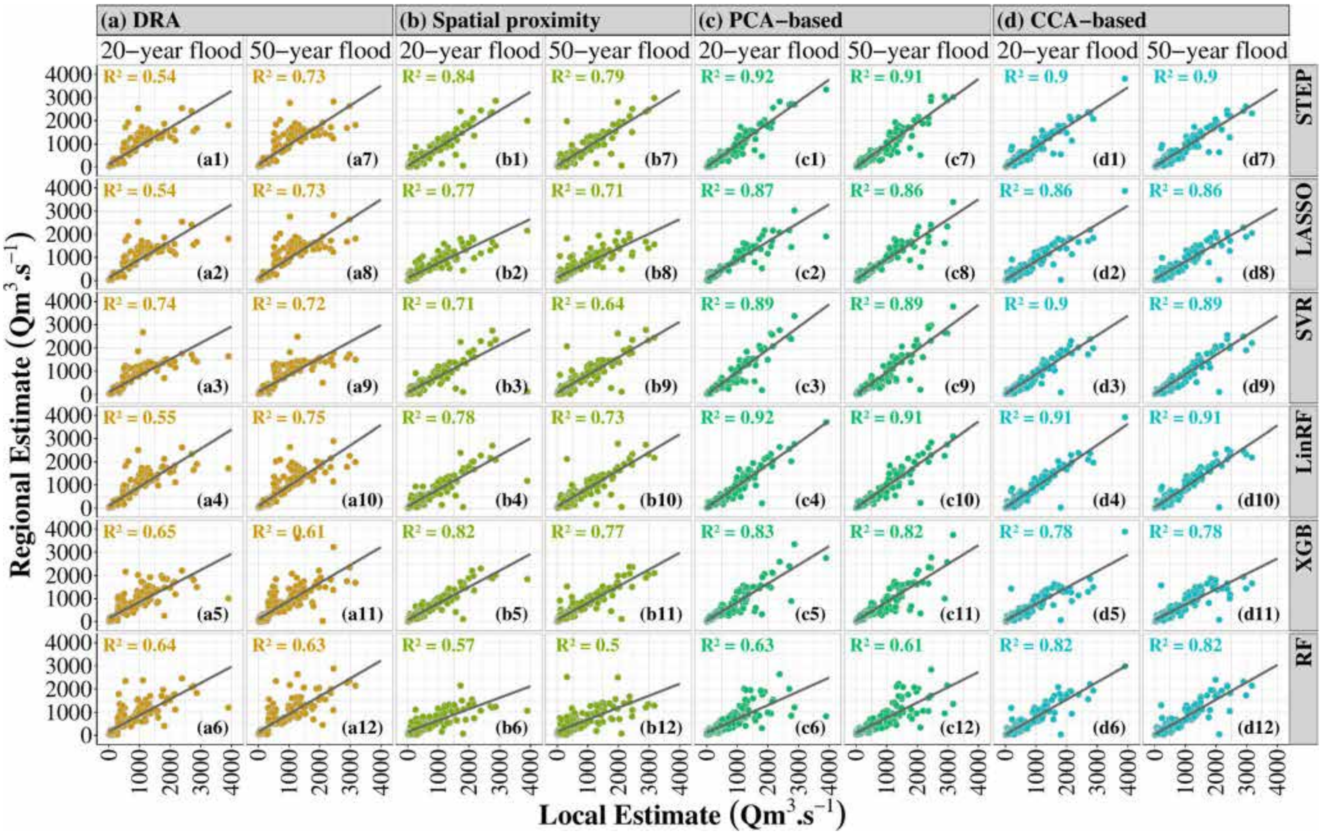

Fig. 5 Regional estimation of 20- and 50-year floods using DRA (**a**) and three approaches for delineation of homogenous regions—Spatial proximity (**b**), PCA (**c**), and CCA (**d**)—applied within the index-flood framework. Both DRA and index-flood approaches are used with two multiple linear regression models (STEP, LASSO), and four ML algorithms (SVR, LinRF, XGB, and RF)

inter-basin variability makes R^2 more sensitive to residuals, which can result in slightly lower values despite smaller absolute errors.

The relative bias (rBias) and MARE metrics for the 20- and 50-year floods, obtained through jackknife sampling across all regional estimation models, are presented in Fig. 6. Overall, all index-flood based methods outperform the DRA. For instance, for the 20-year flood, the DRA shows relatively high MARE (rBias) across the regression models, ranging from 0.47 (0.26) to 0.58 (0.34) for the jackknife sampling (Fig. 5). This suggests that, while the DRA provides a reasonable fit with some regression models, it does not perform as well as regionalization approaches based on the homogeneity assumption. The index-flood framework provides lower systematic errors. Yet, the PCA and CCA approaches outperform the spatial proximity approach. This finding highlights the fact that geographical proximity alone does not ensure homogeneity of watersheds with regard to their hydrological response (Burn 1997). Instead, hydrological proximity should be considered as a function of the physiographic catchment attributes (Ouarda 2016). Figure S2 shows the spatial distribution of at-site flood quantiles. For a consistent spatial pattern analysis, the flood quantiles are scaled with the basin area, thereby minimizing scale effects related to catchment size. In line with the high regional climatic diversity (Emetere 2016; Gbode et al. 2023; Ilori and Ajayi 2020; Muthoni 2020; Vintrou 2012), there is no distinct spatial pattern of flood quantiles distribution across West Africa.

The performances of both PCA and CCA methods vary depending on the regression model used to predict the index-flood, highlighting the importance of model selection in achieving optimal results. The linear models (STEP, LASSO, and SVR) and the mixed linear-tree model (LinRF) show the highest overall performance when combined with the PCA and the CCA frameworks (Fig. 6). Notably, the CCA yields relatively good results when combined with any of the regression and ML models (Fig. 6). Yet, the CCA-SVR combination is the best-suited regional transfer model, yielding the lowest regional estimation errors, with MARE (rBias) values of 0.21 (–0.03) and 0.22 (–0.03), for the 20-year and 50-year floods, respectively (Fig. 6). These findings align with the inter-comparison study by GREHYS (1996a, b), which demonstrated that (i) the neighborhood or "site-focused" approach generally outperforms the "fixed set of regions" method, and (ii) the CCA-based RFFA

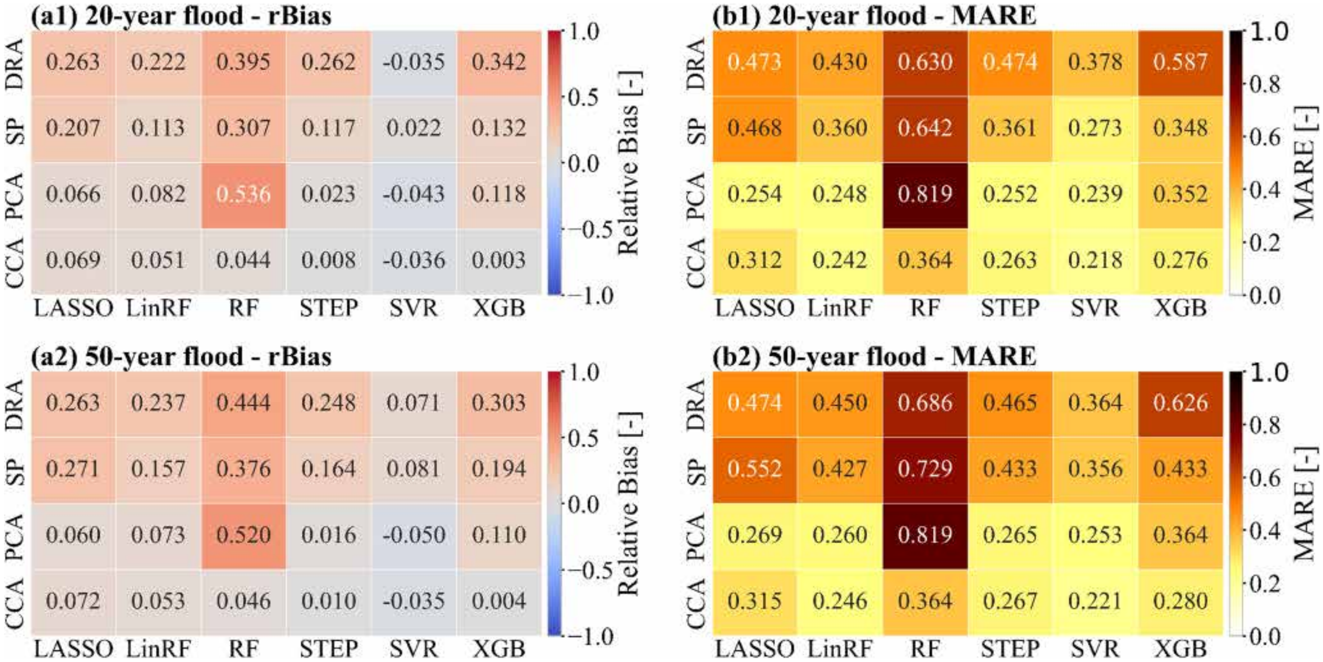

Fig. 6 Relative bias (rBias; panels a1–a2) and mean absolute relative error (MARE; panels b1–b2) for regional flood quantile estimates corresponding to 20-year (top row) and 50-year (bottom row) return periods. All error metrics are computed by comparing regional flood quantile estimates to corresponding local (at-site) values, using a jackknife procedure. Each cell represents the performance of a specific model configuration, combining either the global benchmark (DRA) or one of the index-flood-based regionalization approaches (SP, PCA, or CCA) with a regression/ML algorithm (STEP, LASSO, SVR, LinRF, XGB, and RF)

produces robust results when coupled with various regional estimation procedures. Similarly, comparing four methods for delineating homogeneous regions for a Mexican case study, Ouarda et al. (2008a, b) underscored the advantages of neighborhood-based approaches and demonstrated the superiority of CCA-based methods.

4.3.2 Cross-validation analysis of the regression and ML models

Figure 7 shows the cross-validation (K-fold) results performed within the DRA framework for the regional estimation of 5-, 20-, and 50-year floods. The SVR model consistently outperforms other models' performance, achieving lower bias and error metrics across validation folds and return periods, with MARE (rBias) of 0.34 (0.10), 0.32 (0.05), and 0.33 (0.07), for the 5-, 20- and 50-year floods, respectively (Fig. 7). The high performance provided here by the SVR model aligns with the previous studies of RFFA in various regions (Gizaw and Gan 2016; He et al. 2014; Sharifi Garmdareh et al. 2018; Wu et al. 2008). The robustness of the SVR model with a linear kernel may be partly attributed to the fact that, unlike arid regions, humid catchments typically exhibit more linear hydrological behavior, thereby enhancing predictability (Salinas et al. 2013). Figure S3 shows that most sites within the study area have an aridity index greater than 0.5, and low to moderate prediction errors are associated with higher aridity index values. Moreover, a correlation analysis between the aridity index and the MARE from the SVR model showed a negative relationship (yet not significant), meaning that MARE decreases with humidity. Interestingly, from Figs. 6 and 7, the RF model consistently shows the poorest performance across all information transfer approaches, with higher variability across validation folds, and larger estimation errors. This underperformance contrasts with its well-documented predictive prowess (Desai and Ouarda 2021; Esmaeili-Gisavandani et al. 2023; Mangukiya et al. 2025), and may be explained by their tendency to overfit, especially when trained on limited or skewed datasets (Zhang et al. 2019). For instance, Tramblay et al. (2024) applied the DRA to estimate flood quantiles at ungauged catchments, by comparing the STEP and LASSO regressions, and the RF algorithm. Their results indicated that the best performances are achieved with the two multiple linear regressions.

4.3.3 Spatial patterns of regional flood estimation bias

The spatial distribution of relative bias (rBias) for the regional estimation of the 20-year flood quantile is shown in Fig. 8. Figures S4 and S5 display similar information for the 5- and 50-year floods. Although no clear spatial clustering of regional estimation errors is identified, Figs. 8 and S4–S5 highlight significant differences in performance among

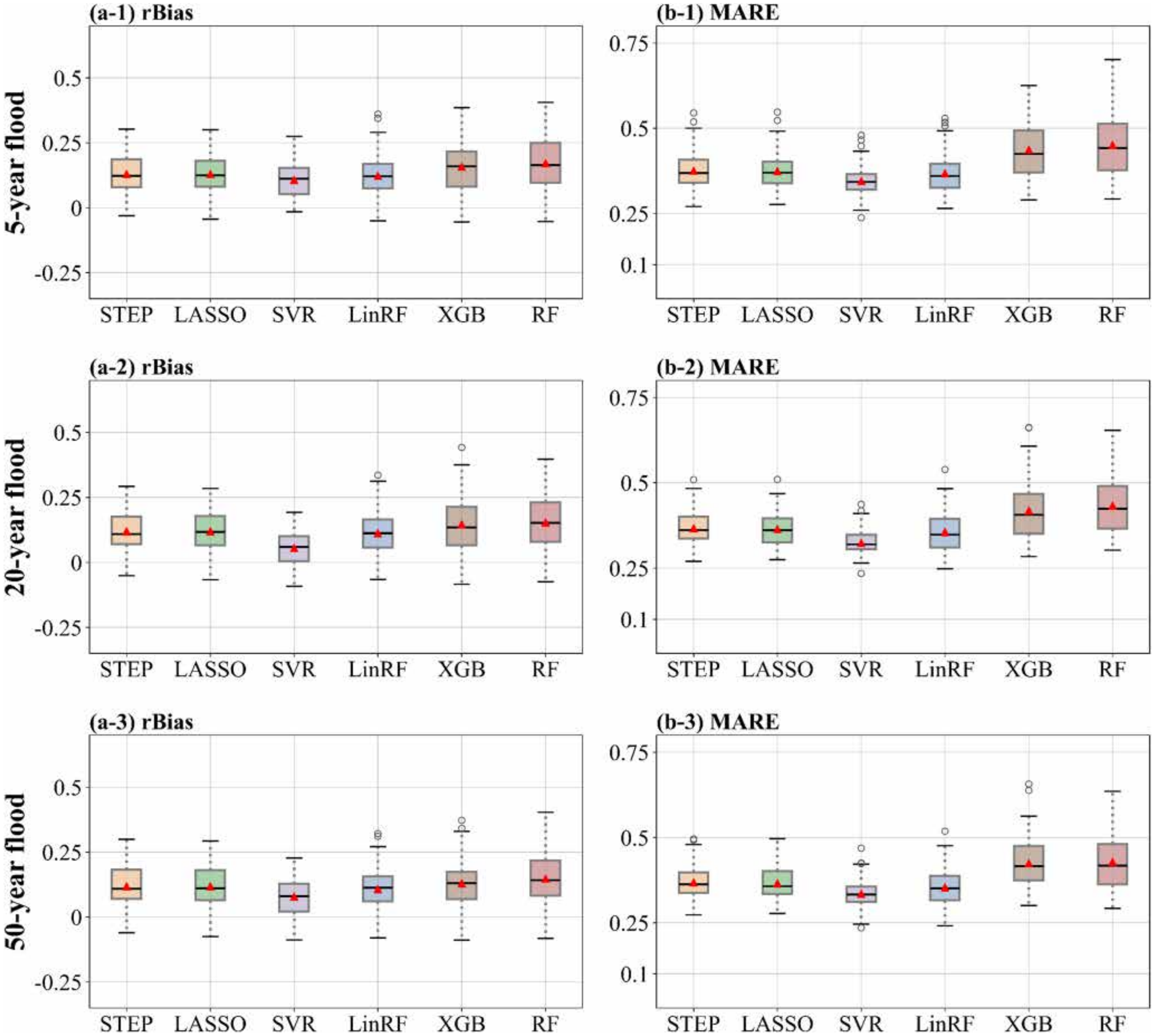

Fig. 7 Cross-validation (K-fold; with K = 100) of the regional estimation of 5-, 20-, and 50-year floods via the DRA, with catchment characteristics as predictors, using different regression and ML models: (i) STEP, (ii) LASSO, (iii) RF, (iv) XGB, (v) SVR, and (vi) LinRF. Performance is assessed using rBias (panels a-1 to a-3) and MARE (panels b-1 to b-3). The red triangles in each boxplot indicate the mean values across the 100 folds

regionalization approaches. The DRA consistently yields pronounced overestimations, reflecting its limited capacity to capture spatial variability in flood-generating processes. The spatial proximity approach reduces this bias moderately but still exhibits extreme overestimations in several catchments. Overall, the CCA method offers the most spatially consistent performance regardless of the regression model, with a more uniform distribution of relative bias and fewer extreme errors. This improvement likely stems from CCA's ability to group catchments based on maximal correlation between physiographic attributes and flood response, resulting in more hydrologically coherent regions. The PCA also yields spatially consistent performance when combined with the SVR model.

4.3.4 Model interpretability and catchment attributes importance

To further assess the individual contribution of predictor variables to flood quantile predictions, we have conducted an additional experiment on relative feature importance. Specifically, we applied the SHapley Additive exPlanations (SHAP; Lundberg and Lee 2017; Shapley 1953) framework to the best-performing regression model, i.e., the

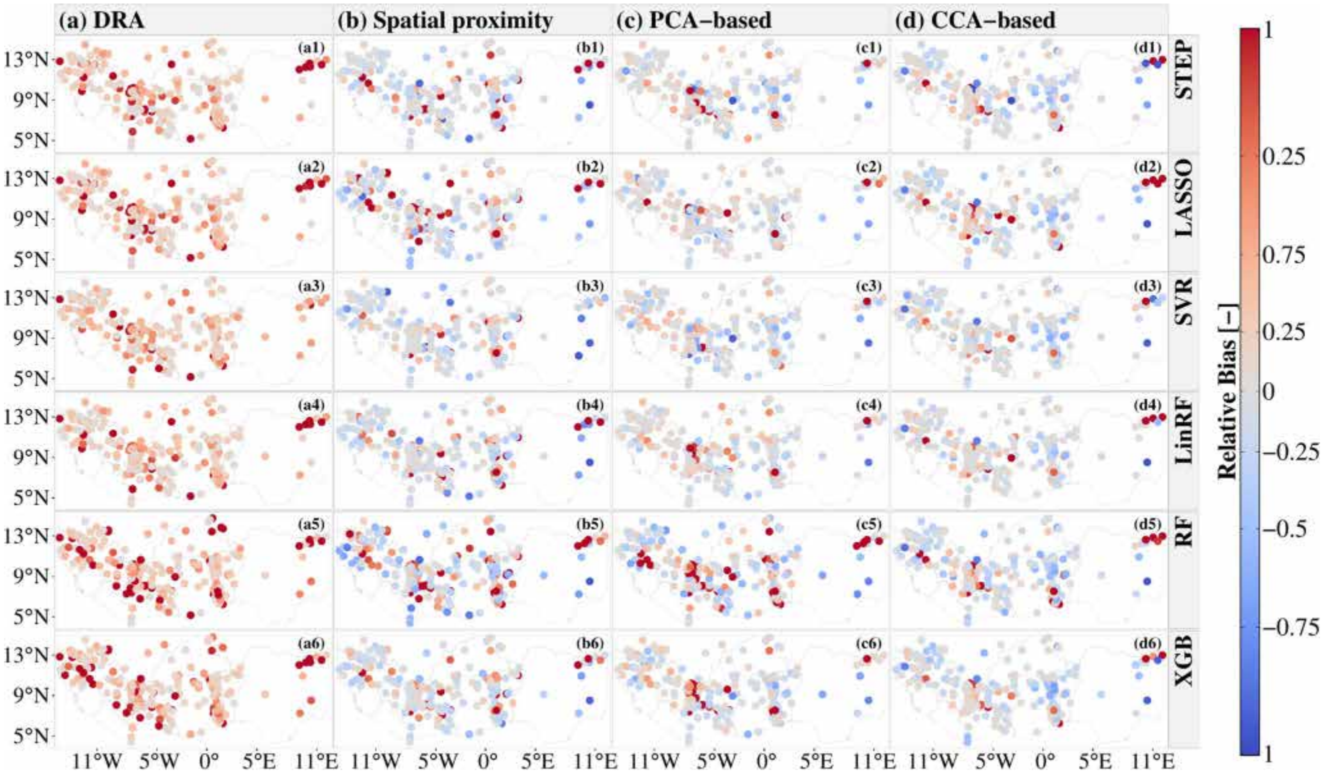

Fig. 8 Spatial distribution of relative bias for the regional estimation of the 20-year flood quantile across West Africa. Columns compare the global benchmark model (DRA) with the three index-flood-based regional delineation approaches (Spatial proximity, PCA-based, and CCA-based). Rows correspond to the different regression and ML models (STEP, LASSO, SVR, LinRF, RF, and XGB). Each point represents a gauging station, and colors indicate the relative bias (dimensionless), with red denoting overestimation and blue denoting underestimation

SVR model. SHAP assigns to each feature an importance value based on its marginal contribution across all possible combinations of input variables, providing a consistent and interpretable approach to quantify how each predictor influences the model's output (He et al. 2023). SHAP values are model-agnostic, meaning they can be used to interpret any machine learning model regardless of its structure. They behave similarly to coefficients in a linear regression model but, unlike other, computational methods like permutation feature importance (Breiman 2001), they capture not only the magnitude of a feature's effect but also its direction (i.e., whether it increases or decreases the predicted flood quantile). The results indicate that the model's predictions are primarily driven by basin geology, topography and scale rather than surface characteristics or rainfall patterns (Fig. 9). For instance, groundwater storage (GS), basin area (Area), groundwater depth (GD), available water capacity (AWC) and maximum altitude (MA) rank as the top-five most influential features (Fig. 9). In addition, high values of all important predictors are associated with higher flood predictions, except for the GD variable (Fig. 8). This supports the hydrological principle that shallow groundwater may limit the soil's ability to absorb rainfall, leading to quicker saturation and higher surface runoff (Teegavarapu and Chinatalapudi 2018). Catchment attributes, such as GP and Altitude, show concentrated distributions of SHAP values, but still meaningful contributions to flood quartile estimations (Fig. 9). The minimal-impact of land use variables (Cropland, Urban, and Shrubland) as shown in Fig. 9a, and their low feature importance values (Fig. 9b) suggests limited predictive power of these variables.

A correlation analysis between model prediction errors and observed input features revealed weak relationships, with no statistically significant correlation coefficients (not shown). Yet, K-means clustering was applied to the prediction errors to explore potential patterns. This unsupervised classification aims to uncover whether specific physiographic characteristics are associated with varying levels of model error. Figure S3 presents the distribution of the most influential predictors across error-based clusters. Catchments associated with higher prediction errors tend to be smaller in area, with lower groundwater storage (GS), shallower groundwater depth (GD), and reduced groundwater productivity (GP). In contrast, catchments with lower errors are generally larger, feature deeper groundwater tables, and more productive aquifers (Figure S3). These catchments are

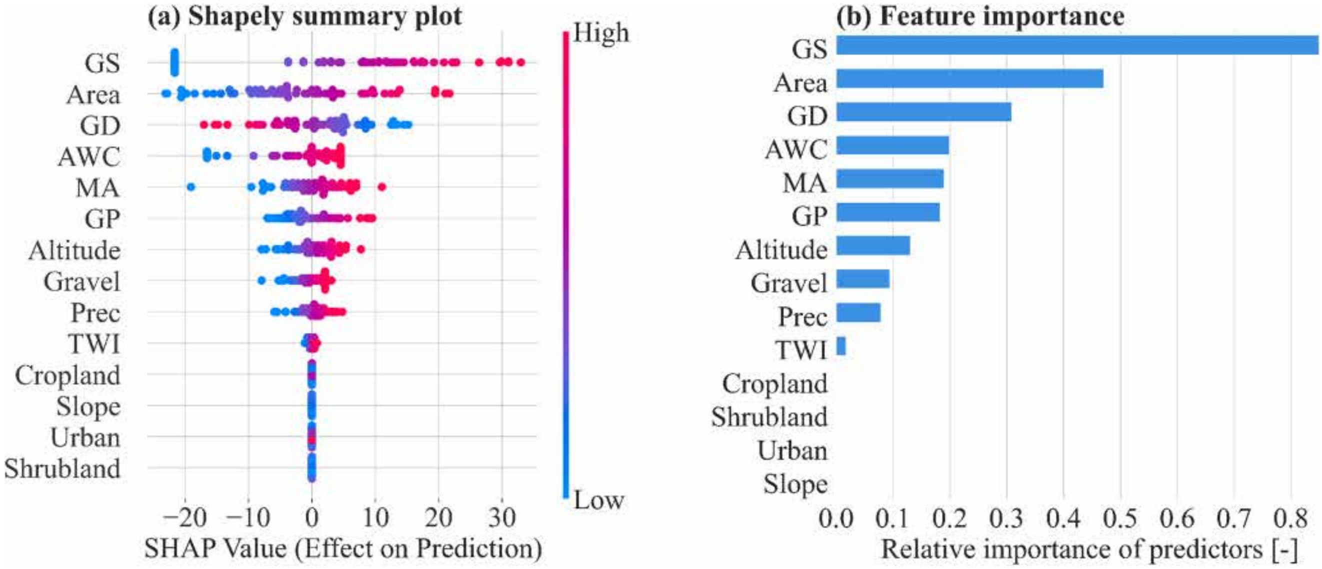

Fig. 9 SHAP (SHapley Additive exPlanations) analysis of the SVR model used to predict the 20-year flood quantile (Q20) across West African catchments. **a** SHAP summary plot showing the distribution of feature effects across the test dataset: the y-axis lists feature names ranked by importance (top to bottom), while the x-axis shows SHAP values representing each feature's contribution to the model output (i.e., impact on predicted Q20). Each dot corresponds to a single observation; color reflects the magnitude of the original feature value, helping to interpret whether high or low values of the feature increase or decrease the prediction. **b** Relative importance of each feature in the dataset

often situated in wetter or humid climates and higher elevation zones. Salinas et al. (2013) stated that the likelihood of finding gauged sub-catchments to serve as donor stations increases with catchment size, thereby improving the reliability of flood characteristic transfers. Moreover, larger catchments tend to exhibit aggregation effects on flood-generating processes, resulting in less flashy flood responses that are generally easier to predict. The study of Salinas et al. (2013) also indicated that prediction error increases with increasing aridity and catchment elevation. Nevertheless, these observed patterns may also reflect imbalances in the dataset, potentially limiting the model's generalization performance in underrepresented regions.

5 Conclusions

This study is the first comprehensive regional flood frequency analysis covering the entire West African region, addressing knowledge gaps in flood risk assessment across a climatically diverse and data-scarce environment. We assessed the performance of a global benchmark model (DRA) and three homogeneous region delineation methods implemented within the index-flood framework (Spatial proximity, PCA, and CCA), each combined with different regression and ML models (STEP, LASSO, SVR, LinRF, RF, and XGB). Under the DRA, models directly estimate flood quantiles, whereas within the index-flood framework they are used to estimate the scale factor. The comparative assessment revealed that index-flood-based approaches consistently outperformed the global DRA benchmark. Among the index-flood based approaches, the framework utilizing CCA demonstrated superior accuracy, while the CCA-SVR combination achieved the lowest regional estimation errors. The robustness of CCA in RFFA has been demonstrated in numerous studies worldwide, including in Canada (Desai and Ouarda 2021; GREHYS 1996a, b; Ouarda et al. 2001; Shu and Ouarda 2007), and Iran (Mohammadi et al. 2017). Feature importance analysis highlighted catchment area and subsurface characteristics as the most influential catchment attributes for regional flood estimation, exerting greater influence than surface features or land use. This finding highlights the critical role of hydrogeological processes in West African flood generation mechanisms. A clustering analysis of prediction errors revealed two distinct patterns: (i) catchments exhibiting higher systematic bias tend to be smaller, with lower groundwater storage and productivity, and shallower water tables; (ii) catchments with lower errors are generally larger, feature deeper groundwater tables, more productive aquifers, and are often located in wetter, higher-elevation regions. These patterns align with the comparative assessment of flood quantile prediction methods at ungauged basins by Salinas et al. (2013) which showed that flood quantile predictions tend to be less accurate in arid than in humid climates, and more accurate in large catchments compared to small ones.

Overall, this study thus demonstrates that robust RFFA is achievable in West Africa through careful methodological design and consideration of dominant physical processes, despite significant data limitations and environmental complexity. As West African communities are facing increasing flood hazards under climate change, these results provide practical pathways to enhance regional resilience through improved risk assessment and infrastructure planning. Yet, despite the comprehensive scope and promising results of this study, it focused primarily on statistical regionalization approaches. These methods rely on the aggregation of hydrological information at gauged sites, which can lead to information loss and reduction of prediction accuracy (Requena et al. 2017). A promising approach is the regional streamflow-based frequency analysis (RSBFA) proposed by Shu and Ouarda (2012) and Requena et al. (2017, 2018). The RSBFA approach first estimates daily streamflow at the ungauged site using a regional flow duration curve (FDC) derived from gauged sites, followed by a local FFA on the extracted extreme streamflow series.

This study advances RFFA in data-scarce and climatically diverse regions, by providing a systematic evaluation of hydrological information transfer approaches across West Africa. The findings contribute to the scientific understanding of catchment-scale flood drivers, and offer a reproducible framework for improving flood estimates at ungauged basins. Beyond methodological insights, the results support more reliable flood risk assessments and infrastructure planning in data-scarce regions increasingly exposed to hydrological extremes under climate change.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00477-026-03258-3.

Acknowledgements The PhD Grant of Serigne Bassirou Diop is funded by the AFD/IRD project CECC. The authors also extend their thanks to the various basin agencies in West Africa for their contribution to data collection and Nathalie Rouche (SIEREM) for the database management.

Author contributions Serigne Bassirou Diop: Conceptualization, Methodology, Data Curation, Writing – Original Draft, Visualization, Investigation, Validation, Writing – Review & Editing. Yves Tramblay: Conceptualization, Methodology, Data Curation, Writing – Original Draft, Validation, Writing – Review & Editing. Ansoumana Bodian: Conceptualization, Methodology, Writing – Review & Editing. Bastien Dieppois: Conceptualization, Methodology, Writing – Review & Editing. Taha B.M.J. Ouarda: Conceptualization, Methodology, Validation, Writing – Review & Editing.

Funding The project was funded by the AFD/IRD project CECC.

Data availability The ADHI dataset containing the observed annual maximum flow time series used in this study is available at https://doi.org/10.23708/LXGXQ9, with Creative Commons Attribution 4.0 International License. The implementation of the codes used in this study relies on the R extRemes library (https://www.jstatsoft.org/article/view/v072i08), and the scikit-learn Python library (https://scikit-learn.org/stable/).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Agoungbome SMD, Seidou O, Thiam M (2018) Evaluation and update of two regional methods (ORSTOM and CIEH) for estimations of flow used in structural design in West Africa. In: Kebe CMF, Gueye A, Ndiaye A, Garba A (eds) Innovations and interdisciplinary solutions for underserved areas. Springer International Publishing, Cham, pp 153–162. https://doi.org/10.1007/978-3-319-98878-8_15

Ahani A, Mousavi Nadoushani SS, Moridi A (2018) A feature weighting and selection method for improving the homogeneity of regions in regionalization of watersheds. Hydrol Process 32(13):2084–2095. https://doi.org/10.1002/hyp.13139

Ahmed A, Yildirim G, Haddad K, Rahman A (2023) Regional flood frequency analysis: a bibliometric overview. Water 15(9):1658. https://doi.org/10.3390/w15091658

Aich V, Liersch S, Vetter T, Fournet S, Andersson JCM, Calmanti S et al (2016) Flood projections within the Niger River Basin under future land use and climate change. Sci Total Environ 562:666–677. https://doi.org/10.1016/j.scitotenv.2016.04.021

Amani A, Paturel J-E (2017) Le projet de révision des normes hydrologiques en Afrique de l'Ouest et Afrique Centrale. La Météorologie 966. https://doi.org/10.4267/2042/61964

Angnuureng BD, Ofosu GT, Ayimbisa EA, Almar R (2025) Evaluating the consequences of the 2023 Akosombo Dam Spillage into the Volta River on Fuveme Beach, Ghana (West Africa). J Coastal Res 113(SI):1021–1025. https://doi.org/10.2112/JCR-SI113-200.1

Ao Y, Li H, Zhu L, Ali S, Yang Z (2019) The linear random forest algorithm and its advantages in machine learning assisted logging regression modeling. J Petrol Sci Eng 174:776–789. https://doi.org/10.1016/j.petrol.2018.11.067

Arnell NW, Gosling SN (2016) The impacts of climate change on river flood risk at the global scale. Clim Change 134(3):387–401. https://doi.org/10.1007/s10584-014-1084-5

Baddianaah I (2023) We all share the blame: analyzing the root causes of flooding in African cities with specific reference to Harper City, Liberia. Environ Challenges 13:100790. https://doi.org/10.1016/j.envc.2023.100790

Baidya S, Singh A, Panda SN (2020) Flood frequency analysis. Nat Hazards 100(3):1137–1158. https://doi.org/10.1007/s11069-019-03853-4

Beijing News (2018) Flood caused by torrential rain in Hami, Xinjiang: 20 people were killed due to partial dam failure of the reservoir. Retrieved from http://news.sina.com.cn/c/2018-08-04/doc-ihhehtqh7551081.shtml

Benito G, O'Connor JE (2013) 9.24 Quantitative paleoflood hydrology. In: Shroder JF (ed) Treatise on geomorphology. Academic, San Diego, pp 459–474. https://doi.org/10.1016/B978-0-12-374739-6.00250-5

Breiman L (2001) Random forests. Mach Learn 45(1):5–32.

Burn DH (1988) Delineation of groups for regional flood frequency analysis. J Hydrol 104(1):345–361. https://doi.org/10.1016/0022-1694(88)90174-6

Burn DH (1990) An appraisal of the region of influence approach to flood frequency analysis. Hydrol Sci J 35(2):149–165. https://doi.org/10.1080/02626669009492415
Burn DH (1997) Catchment similarity for regional flood frequency analysis using seasonality measures. J Hydrol 202(1–4):212–230. https://doi.org/10.1016/S0022-1694(97)00068-1
Cai J, Luo J, Wang S, Yang S (2018) Feature selection in machine learning: a new perspective. Neurocomputing 300:70–79. https://doi.org/10.1016/j.neucom.2017.11.077
Carvalho J, Salvatge Sd'EcopatologiadeF, Departament de Medicina i Cirurgia Animals, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain, Santos JPV et al (2018) Department of Biology, Centre for Environmental and Marine Studies, University of Aveiro, Aveiro 3810–193, Portugal, Sanidad y Biotecnología, Instituto de Investigación en Recursos Cinegéticos, Consejo Superior de Investigaciones Científicas – Universidad de Castilla-La Mancha - Junta de Comunidades de Castilla-La Macha, Ciudad Real 13071, Spain, Torres, R. T.,. Tree-Based Methods: Concepts, Uses and Limitations under the Framework of Resource Selection Models. J Environ Inform. https://doi.org/10.3808/jei.201600352
Cavadias GS (1989) Regional flood estimation by canonical correlation. In Proceedings of the 1989 conference of the Canadian Society of Civil Engineers, St. John's, Newfoundland A, Vol. 11A, pp. 212–231
Cavadias GS (1990) The canonical correlation approach to regional flood estimation. In Proceedings of the Ljubljana Symposium on regionalization in hydrology, IAHS Publ, Vol 191
Cerliani M (2021), April 7 Linear Tree: the perfect mix of Linear Model and Decision Tree. Retrieved March 24, 2025, from https://towardsdatascience.com/linear-tree-the-perfect-mix-of-linear-model-and-decision-tree-2eaed21936b7/
Chagnaud G, Panthou G, Vischel T, Lebel T (2023) Capturing and attributing the rainfall regime intensification in the West African Sahel with CMIP6 models. J Clim 36(6):1823–1843. https://doi.org/10.1175/JCLI-D-22-0412.1
Chebana F, Ouarda TBMJ (2009) Index flood–based multivariate regional frequency analysis. Water Resour Res. https://doi.org/10.1029/2008WR007490
Chebana F, Ouarda TBMJ (2021) Multivariate non-stationary hydrological frequency analysis. J Hydrol 593:125907. https://doi.org/10.1016/j.jhydrol.2020.125907
Chen T, Guestrin C (2016) XGBoost: a scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, ACM, San Francisco, pp 785–794 https://doi.org/10.1145/2939672.2939785
Chokmani K, Ouarda TBMJ (2004) Physiographical space-based kriging for regional flood frequency estimation at ungauged sites. Water Resour Res. https://doi.org/10.1029/2003WR002983
Coles S (2001) An introduction to statistical modeling of extreme values. Springer, London Berlin Heidelberg
Dalrymple T (1960) Flood-frequency analyses. Manual of Hydrology: Part 3. https://doi.org/10.3133/wsp1543A
Dawson CW, Abrahart RJ, Shamseldin AY, Wilby RL (2006) Flood estimation at ungauged sites using artificial neural networks. J Hydrol 319(1):391–409 https://doi.org/10.1016/j.jhydrol.2005.07.032
Dawson CW, Abrahart RJ, Shamseldin AY, Wilby RL, See LM (2005) Neural network modelling of the 20-year flood event for catchments across the UK. In: Proceedings 2005 IEEE international joint conference on neural networks, 2005., vol 4, pp 2637–2642. https://doi.org/10.1109/IJCNN.2005.1556319
Delforge D, Wathelet V, Below R, Sofia CL, Tonnelier M, van Loenhout JAF, Speybroeck N (2025) EM-DAT: the Emergency Events Database. Int J Disaster Risk Reduct 124:105509. https://doi.org/10.1016/j.ijdrr.2025.105509
Desai S, Ouarda TBMJ (2021) Regional hydrological frequency analysis at ungauged sites with random forest regression. J Hydrol 594:125861. https://doi.org/10.1016/j.jhydrol.2020.125861
Ding C, Peng H (2005) Minimum redundancy feature selection from microarray gene expression data. J Bioinform Comput Biol 03(02):185–205. https://doi.org/10.1142/S0219720005001004
Diop SB, Tramblay Y, Bodian A, Ekolu J, Rouché N, Dieppois B (2025) Flood frequency analysis in West Africa. J Flood Risk Manag 18(1):e70001. https://doi.org/10.1111/jfr3.70001
Dorabiala O, Aravkin AY, Kutz JN (2024) Ensemble principal component analysis. IEEE Access 12:6663–6671. https://doi.org/10.1109/ACCESS.2024.3350984
Dosio A, Jury MW, Almazroui M, Ashfaq M, Diallo I, Engelbrecht FA et al (2021) Projected future daily characteristics of African precipitation based on global (CMIP5, CMIP6) and regional (CORDEX, CORDEX-CORE) climate models. Clim Dyn 57(11):3135–3158. https://doi.org/10.1007/s00382-021-05859-w
Douglas EM, Vogel RM, Kroll CN (2000) Trends in floods and low flows in the United States: impact of spatial correlation. J Hydrol 240(1):90–105. https://doi.org/10.1016/S0022-1694(00)00336-X
Drucker H, Burges CJC, Kaufman L, Smola AJ, Vapnik V (1996) Support vector regression machines. In Advances in neural information processing systems 9
Ekolu J, Dieppois B, Diop SB, Bodian A, Grimaldi S, Salamon P et al (2025) How could climate change affect the magnitude, duration and frequency of hydrological droughts and floods in West Africa during the 21st century? A storyline approach. J Hydrol. https://doi.org/10.1016/j.jhydrol.2025.133482
Ekolu J, Dieppois B, Sidibe M, Eden JM, Tramblay Y, Villarini G et al (2022) Long-term variability in hydrological droughts and floods in sub-Saharan Africa: new perspectives from a 65-year daily streamflow dataset. J Hydrol 613:128359. https://doi.org/10.1016/j.jhydrol.2022.128359
El Adlouni S, Ouarda TBMJ, Zhang X, Roy R, Bobée B (2007) Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. Water Resour Res. https://doi.org/10.1029/2005WR004545
Elagib NA, Zayed ISA, Saad SAG, Mahmood MI, Basheer M, Fink AH (2021) Debilitating floods in the Sahel are becoming frequent. J Hydrol 599:126362. https://doi.org/10.1016/j.jhydrol.2021.126362
Emetere ME (2016) Investigations on aerosols transport over micro- and macro-scale settings of West Africa. Environ Eng Res 22(1):75–86. https://doi.org/10.4491/eer.2016.080
England Jr JF, Cohn TA, Faber BA, Stedinger JR, Jr WOT, Veilleux AG et al (2019) Guidelines for determining flood flow frequency — Bulletin 17C (No. 4-B5). Techniques and Methods. U.S. Geological Survey. https://doi.org/10.3133/tm4B5
Eslamian S, Sabzevari Y, Farizi E (2025) Sensitivity analysis of flood frequency on choice of data threshold limit. Results Eng 27:105929. https://doi.org/10.1016/j.rineng.2025.105929
Esmaeili-Gisavandani H, Zarei H, Fadaei Tehrani MR (2023) Regional flood frequency analysis using data-driven models (M5, random forest, and ANFIS) and a multivariate regression method in ungauged catchments. Appl Water Sci 13(6):139. https://doi.org/10.1007/s13201-023-01940-3
FAO (Ed.) (1996) Crues et apports: manuel pour l'estimation des crues décennales et des apports annuels pour les petits bassins versants non jaugés de l'Afrique sahélienne et tropicale sèche. FAO, Rome
Feaster TD, Gotvald AJ, Musser JW, Weaver JC, Kolb K, Veilleux AG, Wagner DM (2023) Magnitude and frequency of floods for rural streams in Georgia, South Carolina, and North Carolina, 2017—Results (No. 2023–5006). Scientific Investigations Report. U.S. Geological Survey. https://doi.org/10.3133/sir20235006
Fisher RA, Tippett LHC (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample. Math Proc

Cambridge Philos Soc 24(2):180–190. https://doi.org/10.1017/S0305004100015681

Fitriana FS, Fischer S, Singh SK, Pahlow M (2025) Regional flood frequency analysis for hydrologically diverse regions in New Zealand. Aust J Water Resour 29(2):135–150. https://doi.org/10.1080/13241583.2025.2530280

Franks SW, Kuczera G (2002) Flood frequency analysis: evidence and implications of secular climate variability, New South Wales. Water Resour Res 38(5):20. https://doi.org/10.1029/2001WR000232

Fréchet M (1927) Sur la loi de probabilité de l'écart maximum. Annales de la Société Polonaise de Mathématique, 6

Gbode IE, Babalola TE, Diro GT, Intsiful JD (2023) Assessment of ERA5 and ERA-interim in reproducing mean and extreme climates over West Africa. Adv Atmos Sci 40(4):570–586. https://doi.org/10.1007/s00376-022-2161-8

Gewers FL, Ferreira GR, Arruda HFD, Silva FN, Comin CH, Amancio DR, Costa LDF (2022) Principal component analysis: a natural approach to data exploration. ACM-CSUR 54(4):1–34. https://doi.org/10.1145/3447755

Girard C (2001) Estimation régionale des crues basée sur l'analyse canonique des corrélations. Université du Québec, INRS-Eau. Retrieved from https://espace.inrs.ca/id/eprint/1559/

Gizaw MS, Gan TY (2016) Regional flood frequency analysis using support vector regression under historical and future climate. J Hydrol 538:387–398. https://doi.org/10.1016/j.jhydrol.2016.04.041

GREHYS (1996a) Inter-comparison of regional flood frequency procedures for Canadian rivers. J Hydrol 186:85–103

GREHYS (1996b) Presentation and review of some methods for regional flood frequency analysis. J Hydrol 186(1):63–84. https://doi.org/10.1016/S0022-1694(96)03042-9

Gumbel EJ (1958) Statistics of extremes. Columbia University, Columbia https://doi.org/10.7312/gumb92958

Haddad K, Rahman A (2012) Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework – Quantile Regression vs. Parameter Regression Technique. J Hydrol 430–431:142–161. https://doi.org/10.1016/j.jhydrol.2012.02.012

Haile AT (2011) Regional flood frequency analaysis in Southern Africa. University of Oslo, Department of Geosciences

Hamlet AF, Lettenmaier DP (2007) Effects of 20th century warming and climate variability on flood risk in the western U.S. Water Resour Res. https://doi.org/10.1029/2006WR005099

Han X, Mehrotra R, Sharma A, Rahman A (2022) Incorporating nonstationarity in regional flood frequency analysis procedures to account for climate change impact. J Hydrol 612:128235. https://doi.org/10.1016/j.jhydrol.2022.128235

Han X, Ouarda TBMJ, Rahman A, Haddad K, Mehrotra R, Sharma A (2020) A network approach for delineating homogeneous regions in regional flood frequency analysis. Water Resour Res 56(3):e2019WR025910. https://doi.org/10.1029/2019WR025910

Harris I., Osborn TJ, Jones P, Lister D (2020) Version 4 of the CRU TS Monthly High-Resolution Gridded Multivariate Climate Dataset. Sci. Data 7, 109 . https://doi.org/10.1038/s41597-020-0453-3

Herrmann SM, Brandt M, Rasmussen K, Fensholt R (2020) Accelerating land cover change in West Africa over four decades as population pressure increased. Commun Earth Environ 1(1):53. https://doi.org/10.1038/s43247-020-00053-y

He Z, Wen X, Liu H, Du J (2014) A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. J Hydrol 509:379–386. https://doi.org/10.1016/j.jhydrol.2013.11.054

He Z, Yang Y, Fang R, Zhou S, Zhao W, Bai Y et al (2023) Integration of shapley additive explanations with random forest model for quantitative precipitation estimation of mesoscale convective systems. Front Environ Sci. https://doi.org/10.3389/fenvs.2022.1057081

Hosking JRM, Wallis JR (1986a) Paleoflood hydrology and flood frequency analysis. Water Resour Res 22(4):543–550. https://doi.org/10.1029/WR022i004p00543

Hosking JRM, Wallis JR (1986b) The value of historical data in flood frequency analysis. Water Resour Res 22(11):1606–1612. https://doi.org/10.1029/WR022i011p01606

Hosking JRM, Wallis JR (1993) Some statistics useful in regional frequency analysis. Water Resour Res 29(2):271–281. https://doi.org/10.1029/92WR01980

Hounkpè J, Diekkrüger B, Afouda AA, Sintondji LOC (2019) Land use change increases flood hazard: a multi-modelling approach to assess change in flood characteristics driven by socio-economic land use change scenarios. Nat Hazards 98(3):1021–1050. https://doi.org/10.1007/s11069-018-3557-8

Hu L, Nikolopoulos EI, Marra F, Anagnostou EN (2020) Sensitivity of flood frequency analysis to data record, statistical model, and parameter estimation methods: an evaluation over the contiguous United States. J Flood Risk Manag 13(1):e12580. https://doi.org/10.1111/jfr3.12580

Ilori OW, Ajayi VO (2020) Change detection and trend analysis of future temperature and rainfall over West Africa. Earth Syst Environ 4(3):493–512. https://doi.org/10.1007/s41748-020-00174-6

James G, Witten D, Hastie T, Tibshirani R (2017) An introduction to statistical learning: with applications in R (Corrected at 8th printing 2017). Springer, New York

Jung K, Kim E, Kang B (2019) Estimation of low-flow in south Korean river basins using a canonical correlation analysis and neural network (CCA-NN) based regional frequency analysis. Atmosphere 10(11):695. https://doi.org/10.3390/atmos10110695

Kay AL, Bell VA, Guillod BP, Jones RG, Rudd AC (2018) National-scale analysis of low flow frequency: historical trends and potential future changes. Clim Change 147(3):585–599. https://doi.org/10.1007/s10584-018-2145-y

Kidson R, Richards KS (2005) Flood frequency analysis: assumptions and alternatives. Progress Phys Geography: Earth Environ 29(3):392–410. https://doi.org/10.1191/0309133305pp454ra

Kim Y, Lee M-J (2020) Rapid change detection of flood affected area after collapse of the Laos Xe-Pian Xe-Namnoy Dam using sentinel-1 GRD data. Remote Sensing 12(12):1978. https://doi.org/10.3390/rs12121978

Komi K, Amisigo B, Diekkrüger B, Hountondji F (2016) Regional flood frequency analysis in the Volta River Basin, West Africa. Hydrology 3(1):5. https://doi.org/10.3390/hydrology3010005

Kuhn M, Johnson K (2013) Applied predictive modeling. Springer, New York

Lalou R, Sultan B, Muller B, Ndonky A (2019) Does climate opportunity facilitate smallholder farmers' adaptive capacity in the Sahel? Palgrave Commun 5(1):1–11. https://doi.org/10.1057/s41599-019-0288-8

Lehner B, Grill G (2013) Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. Hydrol. Process., 27:2171-2186. https://doi.org/10.1002/hyp.9740

Lundberg S, Lee S-I (2017) A unified approach to interpreting model predictions (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1705.07874

Luntz A, Brailovsky V (1969) On estimation of characters obtained in statistical pro-cedure of recognition (in Russian). Techicheskaya Kibernetica, 3

MacDonald AM, Bonsor HC, Dochartaigh BÉÓ, Taylor RG (2012) Quantitative Maps of Groundwater Resources in Africa. Environ. Res. Lett, 7 (2), 024009. https://doi.org/10.1088/1748-9326/7/2/024009

Mahé G (2006) Variabilité pluie-débit en Afrique de l'Ouest et Centrale au 20ème siècle: changements hydro-climatiques, occupation du sol et modélisation hydrologique. Université des Sciences et Techniques Montpellier 2, Montpellier, France. Retrieved from https://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers11-03/010043107.pdf

Mangukiya NK, Kumar A, Sharma A (2025) Machine learning-based regional flood frequency analysis of Indian watersheds. CLEAN – Soil Air Water 53(3):2300120. https://doi.org/10.1002/clen.202300120

Martins ES, Stedinger JR (2000) Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. Water Resour Res 36(3):737–744. https://doi.org/10.1029/1999WR900330

Mohammadi M, Malekinezhad H, Dastorani MT (2017) Comparing Regional Analysis methods for estimation of the peak flow in some watersheds of Isfahan–Sirjan and Yazd-Ardakan Basins. J Range Watershed Manag 70(2):515–529. https://doi.org/10.22059/jrwm.2017.112162.796

Morabbi A, Bouziane A, Seidou O, Habitou N, Ouazar D, Ouarda TBMJ et al (2022) A multiple changepoint approach to hydrological regions delineation. J Hydrol 604:127118. https://doi.org/10.1016/j.jhydrol.2021.127118

Msilini A, Masselot P, Ouarda TBMJ (2020) Regional frequency analysis at ungauged sites with multivariate adaptive regression splines. J Hydrometeorol 21(12):2777–2792. https://doi.org/10.1175/JHM-D-19-0213.1

MsiliniA, Ouarda TBMJ, Masselot P (2022) Evaluation of additional physiographical variables characterising drainage network systems in regional frequency analysis: a Quebec watersheds case-study. Stoch Environ Res Risk Assess 36, 331–351. https://doi.org/10.1007/s00477-021-02109-7

Muhamad Bashar NA, Mohd Arif Zainol MRR, Aziz A, Mazlan MSA, Zawawi AZ, Manan MH A, T. S (2025) Comprehensive review on sustainable dam infrastructure: issues and challenges, factors causing dam failure and future direction in a globally changing climate. Pertanika J Sci Technol 33(3):14. https://doi.org/10.47836/pjst.33.3.14

Muirhead RJ (1982) Aspects of multivariate statistical theory. John Wiley, Hoboken, N. J.

Muthoni F (2020) Spatial-temporal trends of rainfall, maximum and minimum temperatures over West Africa. IEEE J Sel Top Appl Earth Observations Remote Sens 13:2960–2973. https://doi.org/10.1109/JSTARS.2020.2997075

Muthukrishnan R, Rohini R (2016) LASSO: a feature selection technique in predictive modeling for machine learning. In 2016 IEEE international conference on advances in computer applications (ICACA), pp 18–20. https://doi.org/10.1109/ICACA.2016.7887916

NERC NERC (1975) Flood studies report (5 Vols). Meteorological Office, London

Niang I, Ruppel OC, Abdrabo M, Ama E, Lennard C, USA J et al (2014) Chapter 22 Africa. In: Climate Change 2014: impacts, adaptation, and vulnerability. part b: regional aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp 1199–1265

Nicholson SE (2018) Climate of the Sahel and West Africa. In: Nicholson SE (ed) Oxford research encyclopedia of climate science. Oxford University Press, Oxford https://doi.org/10.1093/acrefore/9780190228620.013.510

Nka BN, Oudin L, Karambiri H, Paturel JE, Ribstein P (2015) Trends in floods in West Africa: analysis based on 11 catchments in the region. Hydrol Earth Syst Sci 19(11):4707–4719. https://doi.org/10.5194/hess-19-4707-2015

OCHA (2025) February 20 West and Central Africa: Flooding Situation 2024 Overview - as of 10 February 2025 | United Nations Office for the Coordination of Humanitarian Affairs (OCHA). Retrieved June 11, 2025, from https://www.unocha.org/publications/report/chad/west-and-central-africa-flooding-situation-2024-overview-10-february-2025

Ouarda TBMJ (2016) Regional flood frequency modeling. In: Singh VP (ed) Chow's handbook of applied hydrology, 3rd edn. McGraw-Hill, New York. 77.1–77.8

Ouarda TBMJ, Bâ KM, Diaz-Delgado C, Cârsteanu A, Chokmani K, Gingras H et al (2008b) Intercomparison of regional flood frequency estimation methods at ungauged sites for a Mexican case study. J Hydrol 348(1–2):40–58. https://doi.org/10.1016/j.jhydrol.2007.09.031

Ouarda TBMJ, Girard C, Cavadias GS, Bobée B (2001) Regional flood frequency estimation with canonical correlation analysis. J Hydrol 254(1):157–173. https://doi.org/10.1016/S0022-1694(01)00488-7

Ouarda TBMJ, St-Hilaire A, Bobée B (2008a) Synthèse des développements récents en analyse régionale des extrêmes hydrologiques. Revue Des Sci de l'eau 21(2):219–232. https://doi.org/10.7202/018467ar

Pandey GR, Nguyen V-T-V (1999) A comparative study of regression based methods in regional flood frequency analysis. J Hydrol 225(1–2):92–101. https://doi.org/10.1016/S0022-1694(99)00135-3

Panthou G, Vischel T, Lebel T, Quantin G, Pugin A-CF, Blanchet J, Ali A (2013) From pointwise testing to a regional vision: an integrated statistical approach to detect nonstationarity in extreme daily rainfall. Application to the Sahelian region. J Geophys Res: Atmos 118(15):8222–8237. https://doi.org/10.1002/jgrd.50340

Preti F, Forzieri G, Chirico GB (2011) Forest cover influence on regional flood frequency assessment in Mediterranean catchments. Hydrol Earth Syst Sci 15(10):3077–3090. https://doi.org/10.5194/hess-15-3077-2011

Puech C, Chabi-Gonni D (1983) Méthode de calcul des débits de crue décennale pour les petits et moyens bassins versants en Afrique de l'Ouest et Centrale. Bulletin CIEH (Comite Interafricain d'Etudes Hydrauliques)

Rahman A, Charron C, Ouarda TBMJ, Chebana F (2018) Development of regional flood frequency analysis techniques using generalized additive models for Australia. Stoch Env Res Risk Assess 32(1):123–139. https://doi.org/10.1007/s00477-017-1384-1

Ranstam J, Cook JA (2018) LASSO regression. Br J Surg 105(10):1348–1348. https://doi.org/10.1002/bjs.10895

Requena AI, Ouarda TBMJ, Chebana F (2017) Flood frequency analysis at ungauged sites based on regionally estimated streamflows. J Hydrometeorol 18(9):2521–2539. https://doi.org/10.1175/JHM-D-16-0143.1

Requena AI, Ouarda TBMJ, Chebana F (2018) Low-flow frequency analysis at ungauged sites based on regionally estimated streamflows. J Hydrol 563:523–532. https://doi.org/10.1016/j.jhydrol.2018.06.016

Reynard N, Crooks S, Wilby R, Kay A (2004) Climate change and flood frequency in the UK. 39th Defra Flood and Coastal Flood management Conference, York. https://nora.nerc.ac.uk/id/eprint/2976/

Ribeiro-Corréa J, Cavadias GS, Clément B, Rousselle J (1995) Identification of hydrological neighborhoods using canonical correlation analysis. J Hydrol 173(1–4):71–89. https://doi.org/10.1016/0022-1694(95)02719-6

Robson A, Reed D (1999) Flood Estimation Handbook: Statistical Procedures for Flood Frequency Estimation. Institute of Hydrology. ISBN 978-1-906698-03-4. https://www.ceh.ac.uk/sites/default/files/2021-11/Flood-Estimation-Handbook-3-Statistical-Procedures-For-Flood-Frequency-Estimation_Alice-Robson_Duncan-Reed.pdf

Rodier J, Auvray C (1965) Estimation des débits de crues décennales pour les bassins versants de superficie inférieure à 200 km^2 en Afrique Occidentale. CIEH–ORSTOM, Paris

Rosbjerg D, Blöschl G, Burn DH, Castellarin A, Croke B, Di Baldassarre G et al (2013) Prediction of floods in ungauged basins. In: Viglione A, Blöschl G, Savenije H, Sivapalan M, Wagener T (eds) Runoff prediction in ungauged basins: synthesis across processes, places and scales. Cambridge University Press, Cambridge, pp 189–226. https://doi.org/10.1017/CBO9781139235761.012

Ross CW, Prihodko L, Anchang J, Kumar S, Ji W, Hanan NP (2018) HYSOGs250m, Global Gridded Hydrologic Soil Groups for Curve-Number-Based Runoff Modeling. Sci. Data 5 (1), 180091. https://doi.org/10.1038/sdata.2018.91

Roudier P, Sultan B, Quirion P, Berg A (2011) The impact of future climate change on West African crop yields: What does the recent literature say? Glob Environ Change 21(3):1073–1083. https://doi.org/10.1016/j.gloenvcha.2011.04.007

Salinas JL, Laaha G, Rogger M, Parajka J, Viglione A, Sivapalan M, Blöschl G (2013) Comparative assessment of predictions in ungauged basins – Part 2: flood and low flow studies. Hydrol Earth Syst Sci 17(7):2637–2652. https://doi.org/10.5194/hess-17-2637-2013

Schölkopf B, Bartlett PL, Smola AJ, Williamson RC (1998) Shrinking the tube: a new support vector regression algorithm. In Advances in neural information processing systems vol 11, pp 330–336

Shapley LS (1953) A value for n-person games

Sharifi Garmdareh E, Vafakhah,Mehdi,and, Eslamian SS (2018) Regional flood frequency analysis using support vector regression in arid and semi-arid regions of Iran. Hydrol Sci J 63(3):426–440. https://doi.org/10.1080/02626667.2018.1432056

Shimizu K, Yamada T, Yamada TJ (2020) Uncertainty evaluation in hydrological frequency analysis based on confidence interval and prediction interval. Water 12(9):2554. https://doi.org/10.3390/w12092554

Shu C, Ouarda TBMJ (2007) Flood frequency analysis at ungauged sites using artificial neural networks in canonical correlation analysis physiographic space. Water Resour Res. https://doi.org/10.1029/2006WR005142

Shu C, Ouarda TBMJ (2012) Improved methods for daily streamflow estimates at ungauged sites. Water Resour Res 48(2):2011WR011501. https://doi.org/10.1029/2011WR011501

Sidibe M, Dieppois B, Eden J, Mahé G, Paturel J-E, Amoussou E et al (2019) Interannual to Multi-decadal streamflow variability in West and Central Africa: interactions with catchment properties and large-scale climate variability. Glob Planet Change 177:141–156. https://doi.org/10.1016/j.gloplacha.2019.04.003

Sivakumar B, Singh VP (2012) Hydrologic system complexity and nonlinear dynamic concepts for a catchment classification framework. Hydrol Earth Syst Sci 16(11):4119–4131. https://doi.org/10.5194/hess-16-4119-2012

Smith FGF, Alexander DJ (2018) Quantifying the impact of 20-year flood frequency on land cover change type. Am J Civil Eng Archit 6(1):13–23. https://doi.org/10.12691/ajcea-6-1-2

Song J-H, Her Y, Kang M-S (2022) Estimating reservoir inflow and outflow from water level observations using expert knowledge: dealing with an ill-posed water balance equation in reservoir management. Water Resour Res 58(4):e2020WR028183. https://doi.org/10.1029/2020WR028183

Sørensen R, Zinko U, Seibert J (2006) On the calculation of the topographic wetness index: evaluation of different methods based on field observation. Hydrol. Earth Syst. Sci., 10, 101-112. https://doi.org/10.5194/hess-10-101-2006

Stein L, Clark MP, Knoben WJM, Pianosi F, Woods RA (2021) How do climate and catchment attributes influence flood generating processes? A large-sample study for 671 catchments across the contiguous USA. Water Resour Res 57(4):e2020WR028300. https://doi.org/10.1029/2020WR028300

Tarpanelli A, Paris A, Sichangi AW, O`Loughlin F, Papa F (2023) Water resources in Africa: the role of earth observation data and hydrodynamic modeling to derive river discharge. Surv Geophys 44(1):97–122. https://doi.org/10.1007/s10712-022-09744-x

Taylor CM, Belušić D, Guichard F, Parker DJ, Vischel T, Bock O et al (2017) Frequency of extreme Sahelian storms tripled since 1982 in satellite observations. Nature 544(7651):475–478. https://doi.org/10.1038/nature22069

Teegavarapu RSV, Chinatalapudi S (2018) Incorporating influences of shallow groundwater conditions in curve number-based runoff estimation methods. Water Resour Manage 32(13):4313–4327. https://doi.org/10.1007/s11269-018-2053-y

Tibshirani R (1996) Regression shrinkage and selection via the lasso. J R Stat Soc: Ser B (Methodol) 58(1):267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tramblay Y, El Khalki EM, Khedimallah A, Sadaoui M, Benaabidate L, Boulmaiz T et al (2024) Regional flood frequency analysis in North Africa. J Hydrol 630:130678. https://doi.org/10.1016/j.jhydrol.2024.130678

Tramblay Y, Rouché N, Paturel J-E, Mahé G, Boyer J-F, Amoussou E et al (2021b) ADHI: the African Database of Hydrometric Indices (1950–2018). Earth Syst Sci Data 13(4):1547–1560. https://doi.org/10.5194/essd-13-1547-2021

Tramblay Y, Somot S (2018) Future evolution of extreme precipitation in the Mediterranean. Clim Change 151:289–302. https://doi.org/10.1007/s10584-018-2300-5

Tramblay Y, Villarini G, El Khalki EM, Gründemann G, Hughes D (2021a) Evaluation of the drivers responsible for flooding in Africa. Water Resour Res 57(6):e2021WR029595. https://doi.org/10.1029/2021WR029595

Tramblay Y, Villarini G, Zhang W (2020) Observed changes in flood hazard in Africa. Environ Res Lett 15(10):1040b5. https://doi.org/10.1088/1748-9326/abb90b

Vapnik V (1995) The nature of statistical learning theory. Springer, New York

Vintrou E (2012) Cartographie et caractérisation des systèmes agricoles au Mali par télédétection à moyenne résolution spatiale (PhD Thesis). AgroParisTech

Weibull W (1951) A statistical distribution function of wide applicability. J Appl Mech 18(3):293–297. https://doi.org/10.1115/1.4010337

Wieder W (2014) Regridded Harmonized World Soil Database v1.2. 59.234908 MB. https://doi.org/10.3334/ORNLDAAC/1247

Wilcox C, Vischel T, Panthou G, Bodian A, Blanchet J, Descroix L et al (2018) Trends in hydrological extremes in the Senegal and Niger Rivers. J Hydrol 566:531–545. https://doi.org/10.1016/j.jhydrol.2018.07.063

World Bank (2021) World Bank Engagement in transboundary waters in west Africa: retrospective and lessons learned. World Bank, Washington

Wu CL, Chau KW, Li YS (2008) River stage prediction based on a distributed support vector regression. J Hydrol 358(1):96–111. https://doi.org/10.1016/j.jhydrol.2008.05.028

Yang X, Chen H, Wang Y, Xu C-Y (2016) Evaluation of the effect of land use/cover change on flood characteristics using an integrated approach coupling land and flood analysis. Hydrol Res 47(6):1161–1171. https://doi.org/10.2166/nh.2016.108

Younes K, Kharboutly Y, Antar M, Chaouk H, Obeid E, Mouhtady O et al (2023) Application of unsupervised machine learning for the evaluation of aerogels' efficiency towards ion removal—a principal component analysis (PCA) Approach. Gels 9(4):304. https://doi.org/10.3390/gels9040304

Zhang H, Nettleton D, Zhu Z (2019) Regression-enhanced random forests. arXiv. https://doi.org/10.48550/arXiv.1904.10416