REPUBLIC OF BENIN

------000------

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

------000------

UNIVERSITY OF ABOMEY - CALAVI

------000------

DOCTORAL SCHOOL OF LIFE AND EARTH SCIENCES

------000------

Registered under N°: 470

A DISSERTATION

Submitted

In partial fulfillment of the requirements for the degree of

DOCTOR of Philosophy (PhD) of the University of Abomey-Calavi

In the framework of the

Graduate Research Program on Climate Change and Water Resources (GRP-CCWR)

By

Fousseini KOUYATE

Public defense on: 10/15/2025

Subject: Hydrology
Speciality: Climate Change and Water Resources

======

MACHINE LEARNING APPLICATIONS FOR PREDICTING WATER LEVEL FLUCTUATIONS IN THE BANI RIVER BASIN UNDER CLIMATE CHANGE SCENARIOS

======

Supervisor:

Kossi François GUEDJE	Associate Professor	University of Abomey-Calavi, UAC, Benin

Reviewers:

Adéchina Eric ALAMOU	Full Professor	UNSTIM, Benin
Markus DISSE	Full Professor	Technical University of Munich, Germany
Olatunji OLUWASEMIRE	Full Professor	University of Ibadan, Oyo State, Nigeria

JURY

Agnidé Emmanuel LAWIN	Full Professor, University of Abomey Calavi, UAC, Benin	President
Adéchina Eric ALAMOU	Full Professor, UNSTIM, Benin	Reviewer
Markus DISSE	Full Professor, Technical University of Munich, Germany	Reviewer
Olatunji OLUWASEMIRE	Full Professor, University of Ibadan, Oyo State, Nigeria	Reviewer
Aymar BOSSA	Associate Professor, University of Abomey Calavi, UAC, Benin	Examiner
Milad AMINZADEH	Senior scientist, Hamburg University of Technology, Germany	Advisor
Kossi François GUEDJE	Associate Professor, University of Abomey-Calavi, UAC, Benin	Supervisor

REPUBLIQUE DU BENIN

------000------

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

------000------

UNIVERSITE D'ABOMEY - CALAVI

------000------

ECOLE DOCTORALE SCIENCES DE LA VIE ET DE LA TERRE

------000------

Enregistrée sous N°: 470

THESE

Soumise pour obtenir le grade de

DOCTEUR de l'Université d'Abomey-Calavi

Dans la Spécialité:

Changement climatique et Ressources en Eau

Par

Fousseini KOUYATE

Soutenue publiquement le : 15/10/2025

Discipline: Hydrologie
Spécialité: Changements Climatiques et Ressources en Eau

======

APPLICATIONS DE L'APPRENTISSAGE AUTOMATIQUE POUR PREDIRE LES FLUCTUATIONS DU NIVEAU D'EAU DANS LE BASSIN DE LA RIVIERE BANI SOUS LES SCENARIOS DE CHANGEMENT CLIMATIQUE

Directeur de thèse:

Kossi François GUEDJE	Maître de Conférences	Université d'Abomey-Calavi, UAC, Benin

Rapporteurs:

Adéchina Eric ALAMOU	Professeur Titulaire	UNSTIM, Bénin
Markus DISSE	Professeur Titulaire	Université technique de Munich, Allemagne
Olatunji OLUWASEMIRE	Professeur Titulaire	Université d'Ibadan, Oyo State, Nigéria

JURY

Agnidé Emmanuel LAWIN	Professeur Titulaire, Université d'Abomey-Calavi, UAC, Bénin	Président
Adéchina Eric ALAMOU	Professeur Titulaire, UNSTIM, Bénin	Rapporteur
Markus DISSE	Professeur Titulaire, Université technique de Munich, Allemagne	Rapporteur
Olatunji OLUWASEMIRE	Professeur Titulaire, Université d'Ibadan, Oyo State, Nigéria	Rapporteur
Aymar BOSSA	Maître de Conférences, Université d'Abomey-Calavi, UAC, Bénin	Examinateur
Milad AMINZADEH	Scientifique Senior, Université des Technologies de Hambourg, Allemagne	Membre de l'encadrement
Kossi François GUEDJE	Maître de Conférences, Université d'Abomey-Calavi, UAC, Bénin	Directeur de thèse

Dedication

To my beloved Father

To my beloved Mother

Brothers, Sister and Wife

This work is yours

Acknowledgment

This PhD work is realized in the framework of the West African Science Service Center on Climate Change and Adapted Land Use (WASCAL) and funded by the German Ministry of Education and Research (BMBF) in collaboration with the Benin Ministry of Higher Education and Scientific Research (MESRS).

My utmost gratitude always goes to ALLAH Almighty who sustains my life from my birth up to today. I know He will still stand by my side every day of my life.

I want to thank the German Ministry of Education and Research, which gave me, at first hand, through the wonderful scholarship under the framework of WASCAL, the opportunity to undertake doctoral studies in Climate Change and Water Resources at the University of Abomey-Calavi, Benin Republic.

This work could not have been achieved without the involvement of a number of individuals and institutions. This is the opportunity given to me to thank all of them.

I will continue by thanking the Vice Chancellor of the University of Abomey-Calavi, who accepted me into his university; may he find all my gratitude.

I am very grateful to my Supervisor, Prof. Kossi François Guédjé, who has always been available during this work to correct the papers, advise me, and motivate me to finish on time. It has been hard, but now I can see the importance of all the encouragement he has given me. Thank you very much, dear Professor; I will be very honored to continue working under your guidance even after this PhD.

I would like to express deep gratitude to Prof. Nima Shokri and Dr. Milad Aminzadeh from the Institute of Geo-Hydroinformatics at Hamburg University of Technology, Hamburg, Germany. Their advice, support, and daily encouragement throughout my thesis journey have been invaluable. The chance to spend four months of research stay at their esteemed institute greatly enriched my research experience.

I thank all the professors who contributed to my training at the University of Abomey-Calavi, especially during the 6-month coursework. I am very proud to have learned from you. May the Almighty reward you.

Abstract

This study addresses the hydroclimatic challenges of the Bani River Basin (BRB) in southeastern Mali, where data scarcity, climate variability, and socio-political instability severely complicate water resource management. First, we validate satellite-based rainfall products (e.g., CHIRPS) using available ground observations to determine their reliability in capturing rainfall patterns. The results confirm the usefulness of these products, particularly in regions where insecurity and logistical constraints limit in-situ measurements. Second, we employ statistical trend analyses to examine the spatiotemporal evolution of rainfall and river discharge indices over recent decades. The findings reveal shifting extremes more frequent short-duration heavy rainfall and intensifying peak flows, which raises concerns about water scarcity, flood risks, and agricultural stability. Third, we develop and test machine learning (ML) models, including Random Forest and Gradient Boosting, to replicate historical water-level fluctuations. Beyond prediction, these ML models can also be employed to explore the underlying relationships between rainfall and water levels, offering insights into the hydrological response to rainfall variability. These models demonstrate strong predictive capability, outperforming traditional methods when faced with complex, nonlinear datasets. Finally, by integrating the Coupled Model Intercomparison Project Phase 6 (CMIP6) projections under scenarios SSP2-4.5 and SSP5-8.5, we forecast future water levels and identify key breakpoints in hydrological regimes. The results underscore an urgent need for adaptive management strategies, including improved monitoring infrastructures, data-sharing mechanisms, and community-driven interventions. Overall, this research offers actionable insights for researchers, policymakers and stakeholders, emphasizing the viability of satellite-based data and advanced computational methods to mitigate climate risks and strengthen water governance in the BRB.

Keywords: Bani River Basin; Data scarcity; Climate variability; Socio-political instability; Water resource management; Machine learning; Satellite-based rainfall; CMIP6 scenarios

Résumé

Cette étude aborde les défis hydroclimatiques du bassin de la rivière Bani (BRB) dans le sud-est du Mali, où la rareté des données, la variabilité climatique et l'instabilité socio-politique rendent la gestion des ressources en eau particulièrement complexe. Dans un premier temps, nous évaluons la fiabilité des estimations satellitaires de précipitations (p. ex. CHIRPS) en les comparant aux observations au sol disponibles. Les résultats confirment la pertinence de ces produits, surtout dans des zones où l'insécurité et les contraintes logistiques limitent les mesures in situ. Deuxièmement, nous appliquons des analyses statistiques de tendance pour examiner l'évolution spatiotemporelle des précipitations et des débits fluviaux au cours des dernières décennies. Les conclusions révèlent des extrêmes changeants avec des épisodes de pluies intenses de courte durée plus fréquents et des débits de crue plus marqués faisant craindre une accentuation des pénuries d'eau, des risques d'inondation et des impacts sur l'agriculture. Troisièmement, nous développons et testons des modèles d'apprentissage automatique, notamment Random Forest et Gradient Boosting, pour reproduire les fluctuations historiques du niveau de l'eau. Ces modèles montrent une forte capacité de prévision, surpassant les méthodes classiques face à des jeux de données complexes et non linéaires. Enfin, grâce à l'intégration des projections du Coupled Model Intercomparison Project Phase 6 (CMIP6) selon les scénarios SSP2-4.5 et SSP5-8.5, nous prévoyons l'évolution future des niveaux d'eau et identifions des points de rupture clés. Les résultats soulignent l'urgence de stratégies de gestion adaptative, incluant le renforcement des infrastructures de suivi, le partage de données et l'implication des communautés. Dans l'ensemble, cette étude démontre la fiabilité des données satellitaires, l'intensification des extrêmes hydroclimatiques et la performance supérieure des modèles d'apprentissage automatique, soulignant leur potentiel pour anticiper les fluctuations hydrologiques et renforcer la gouvernance de l'eau dans le bassin de la rivière Bani.

Mots-clés : Bassin de la rivière Bani; Rareté des données; Variabilité climatique; Instabilité socio-politique; Gestion des ressources en eau; Apprentissage automatique; Précipitations satellitaires; Scénarios CMIP6

Introduction

Les ressources en eau dans les régions semi-arides du monde sont de plus en plus affectées par la convergence des changements climatiques, des modifications induites par l'homme et des instabilités socio-politiques, ce qui augmente la pression sur ces systèmes vitaux (O'Brien et al., 2020). En Afrique de l'Ouest, ces défis sont particulièrement cruciaux en raison de la dépendance des communautés locales envers les systèmes fluviaux pour des activités telles que l'agriculture et la pêche (Roudier and Mahe, 2010a; Sylla et al., 2018). Le bassin de la rivière Bani (BRB) au Mali, tributaire du fleuve Niger, est vital pour l'agriculture, les écosystèmes de zones humides, et la pêche, bien que les recherches y soient moins développées que dans d'autres bassins comme celui du fleuve Sénégal (Mahe et al., 2013; Traore, 2015).

Les études hydrologiques récentes menées dans le bassin de la rivière Bani mettent en évidence d'importantes lacunes, tant dans la disponibilité des données que dans la précision des modèles prédictifs. Cette insuffisance, aggravée par l'insécurité et les contraintes logistiques limitant la collecte des données in situ, est renforcée par des anomalies climatiques marquées. Elle complique considérablement la gestion des ressources en eau, notamment dans des zones sensibles comme Mopti, où les décisions stratégiques reposent souvent sur des informations partielles ou obsolètes (Andersson et al., 2017; Onuoha et al., 2023).

Pour surmonter ces obstacles, l'utilisation de données satellitaires, comme celles fournies par la Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), a commencé à émerger comme une solution viable. Ces produits, qui combinent télédétection et données au sol, offrent une couverture plus complète que les réseaux conventionnels et sont essentiels dans les régions affectées par les conflits (Houngnibo et al., 2023a). Cependant, leur validation reste essentielle pour garantir leur fiabilité, surtout dans les zones où les données au sol sont insuffisantes (Dinku et al., 2010).

En parallèle, les avancées dans les techniques d'apprentissage automatique (machine learning, ML), telles que les modèles Random Forest (RF), Gradient Boosting (GB), Extreme Gradient Boosting (XGB) et Support Vector Machine (SVM), offrent de nouvelles possibilités pour la modélisation et la prévision des phénomènes hydrologiques. Ces modèles peuvent traiter des

données complexes et hétérogènes, rendant les prévisions plus précises, notamment en intégrant des scénarios climatiques futurs du CMIP6 pour anticiper les impacts des changements climatiques (Hassani et al., 2024; Tahmasebi and Sahimi, 2021).

C'est dans ce contexte que la présente thèse ambitionne de renforcer la surveillance et la prévision hydrologiques pour le BRB en s'appuyant sur quatre objectifs fondamentaux :

- Évaluer la fiabilité des produits satellitaires afin de combler les lacunes en termes de données.
- Analyser les tendances et les schémas hydroclimatiques sur les dernières décennies.
- Évaluer les performances prédictives de quelques modèles de ML dans la reproduction des niveaux d'eau historiques.
- Prévoir les niveaux d'eau futurs selon les scénarios climatiques SSP2-4.5 et SSP5-8.5.

En définitive, cette recherche entend enrichir les connaissances sur le BRB et proposer des stratégies de gestion adaptées aux différentes trajectoires d'émission de gaz à effet de serre, contribuant significativement à l'amélioration de la gestion des ressources hydriques dans une région complexe et variable (O'Brien et al., 2020; Sawadogo et al., 2024).

Zone d'études

Cette étude porte sur le BRB, dont l'exutoire se situe à Mopti, au centre du Mali. Le bassin transfrontalier de la Bani s'étend sur une superficie d'environ 129 400 km², entre les latitudes 9°10'13"N et 14°05'44"N et les longitudes 8°05'16"O et 3°46'17"O. Bien que la majeure partie du bassin se trouve au Mali, une partie s'étend en Côte d'Ivoire et au Burkina Faso (Chaibou Begou et al., 2016). Le relief du bassin versant est peu élevé, avec une pente douce allant de 852 m dans les régions sud et centre-est à 217 m à l'exutoire nord. Les terres agricoles constituent la principale pratique d'occupation des sols, aux côtés de vastes étendues de savane et de zones forestières (Loveland et al., 2000). Le BRB est soumis à un climat soudano-sahélien ; la rivière s'écoule du sud vers le nord, traversant un important gradient pluviométrique. Les données historiques indiquent une précipitation annuelle moyenne allant de 1250 mm à Odienné à 615 mm à Ségou, selon les moyennes de 1981 à 2000. Le débit annuel

moyen enregistré à la station de Douna pendant la même période s'établissait à environ 184 m³/s, soit une lame d'eau de ruissellement de 58 mm pour une pluviométrie annuelle moyenne de 1000 mm (Amogu et al., 2010). Des valeurs de ruissellement particulièrement faibles ont été relevées en 1983, 1984 et 1987. La région a subi une forte baisse pluviométrique à partir de 1970-1971, qui a perduré durant deux décennies et profondément affecté les ressources en eau de ce bassin.

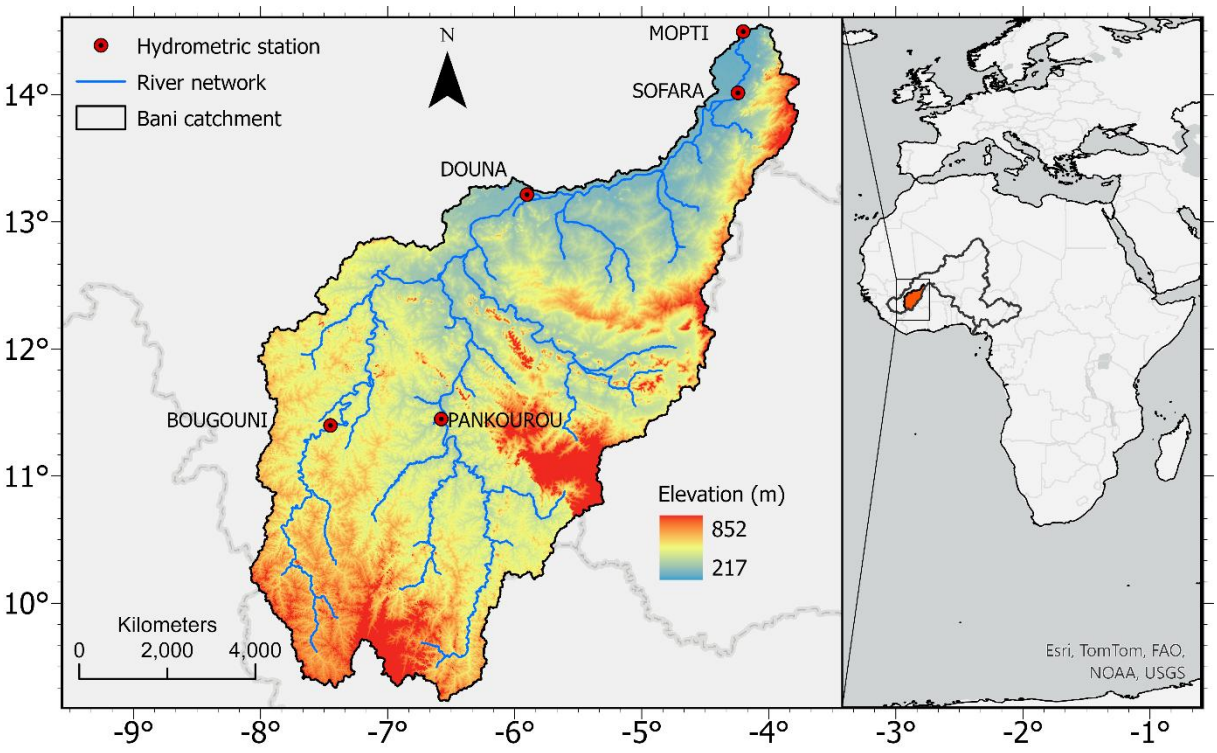

Figure 1: BRB en Afrique de l'ouest, partagé entre le Mali, la Côte d'Ivoire et le Burkina Faso.

Données et méthodes

Cette étude s'appuie sur un large ensemble de données hydrologiques et météorologiques collectées depuis 1983 afin de proposer un cadre de prévision des fluctuations du niveau d'eau dans le BRB. Cinq stations de mesure (Bougouni, Pankourou, Douna, Sofara et Mopti) fournissent des données journalières de hauteur d'eau, bien que l'insécurité récurrente dans la région de Mopti limite la continuité des observations à cette station. Parallèlement, des relevés quotidiens de précipitations, d'évaporation totale et de profondeur de nappe, ainsi que des estimations satellitaires (CHIRPS, PERSIANN-CDR), ont été intégrés pour renforcer la

robustesse du modèle, compte tenu de la rareté des stations pluviométriques locales (Oyerinde et al., 2017). Pour remédier à l'insuffisance des données in-situ, deux produits de précipitations satellitaires ont été évalués : CHIRPS (Funk et al., 2015) et PERSIANN-CDR (Ashouri et al., 2015). Dans ce contexte, des corrections de biais sont nécessaires afin de pallier les incertitudes inhérentes à ces données, particulièrement en zones peu instrumentées. De plus, le modèle s'appuie sur des projections climatiques issues de dix simulations CMIP6, couvrant la période 1983–2050 pour les températures et précipitations quotidiennes. Deux scénarios d'émission ont été considérés, SSP2-4.5 et SSP5-8.5, afin de couvrir un large éventail d'évolutions potentielles des gaz à effet de serre (Limantol et al., 2023; Sawadogo et al., 2024). Cette approche vise à tenir compte de la vulnérabilité marquée de l'Afrique de l'Ouest aux événements hydroclimatiques extrêmes, tout en améliorant la qualité et la fiabilité des prévisions pour le BRB.

Cette recherche repose sur un cadre méthodologique conçu pour atteindre ses objectifs spécifiques, en combinant des techniques d'analyse de données avec des approches d'interprétation rigoureuses.Il commence par la section 3.1, qui traite de la collecte, du prétraitement et de l'intégration des données historiques de précipitations issues de sources satellitaires (PERSIANN-CDR et CHIRPS) et in situ. Un Produit de Référence des Précipitations fiable est alors développé et validé à travers des techniques géostatistiques, facilitant une évaluation robuste des estimations dérivées des satellites. Dans la section 3.2, l'accent est mis sur l'identification et la caractérisation des événements hydroclimatiques extrêmes, incluant les indices de précipitations et de débits fluviaux, qui sont essentiels pour comprendre les risques d'inondation et de sécheresse. Cette section détaille également les méthodes de détection de tendances et de points de changement (Mann-Kendall Modifié, test de Pettitt et SNHT) utilisées pour discerner les motifs à long terme et les changements abrupts dans le régime hydrologique du bassin.

La section 3.3 introduit l'implémentation de quatre algorithmes d'apprentissage automatique : Extreme Gradient Boosting (XGB), Support Vector Machine (SVM), RF et GB, pour la prévision des fluctuations du niveau de l'eau à l'exutoire du bassin versant Bani. Un schéma de

validation croisée assure que les modèles prédictifs sont testés sur plusieurs segments temporels pour améliorer leur capacité à bien se comporter sur des données nouvelles et leur fiabilité. La section 3.4 discute des projections hydroclimatiques futures dérivées de différents scénarios climatiques CMIP6 (SSP2-4.5 et SSP5-8.5). Les résultats de précipitations et de températures issus d'une série de modèles globaux sont analysés pour prévoir les changements dans des indicateurs clés tels que les jours consécutifs secs/humides, la précipitation journalière maximale et les niveaux d'eau de crue. Cette approche intégrée, combinant données historiques, interpolation, apprentissage automatique et modélisation des scénarios climatiques, offre une compréhension complète des conditions hydroclimatiques passées, présentes et futures du bassin.

Résultats

Le BRB, situé en Afrique de l'Ouest, joue un rôle essentiel en tant qu'affluent du fleuve Niger (Dos Santos, 2023). Ses ressources en eau soutiennent l'agriculture, la pêche et l'approvisionnement en eau pour plusieurs millions de personnes (Roudier and Mahe, 2010a). Toutefois, ce bassin demeure moins étudié que les grands bassins fluviaux d'Afrique de l'Ouest comme le Sénégal ou le Niger et fait face à des défis croissants liés aux changements climatiques, notamment des sécheresses et des précipitations extrêmes, ainsi qu'à des pressions anthropiques telles que la construction de barrages et l'évolution de l'occupation des sols (Diop et al., 2020; Sylla et al., 2018). Afin de mieux caractériser la variabilité spatiale de la pluviométrie, une analyse variographique exponentielle a été réalisée sur des données observées, révélant une portée d'environ 1,8° (≈ 200 km) et un sill à 78 km. Cette information a servi de base à l'élaboration d'un produit de référence (RRP) par krigeage ordinaire à une résolution de 0,25°. Les indicateurs statistiques (MAE, NRMSE et PBIAS) montrent que CHIRPS présente globalement des erreurs plus faibles et un biais plus neutre que PERSIANN-CDR. Par exemple, en 1989, CHIRPS affichait un MAE de 3,23 mm, un NRMSE de 94,1 et un PBIAS de -30,7 %, alors que PERSIANN-CDR enregistrait un MAE de 4,01 mm, un NRMSE de 100,6 et un PBIAS de +8,6 %, ce qui indique une tendance à la surestimation pour PERSIANN-CDR et une légère sous-estimation pour CHIRPS. L'étude des indices de

précipitations extrêmes (RX1DAY, RX5DAY, R95P, R99P, SDII) sur la période 1991–2020 met en évidence une plus forte intensité des pluies extrêmes dans la partie sud du bassin, tandis que le nord-est se caractérise par des hausses significatives de certains indices ($p < 0,05$), notamment R99P et RX1DAY. Des tests de tendance (Mann–Kendall modifié) et de détection de ruptures (Pettitt, SNHT) indiquent des points de rupture autour de 2000 pour RX5DAY et de 2012 pour RX1DAY, soulignant la complexité du régime pluviométrique et la sensibilité aux changements climatiques déjà observés dans d'autres bassins ouest-africains (Ndiaye et al., 2023; Roudier and Mahe, 2010a). L'analyse des débits extrêmes (Q95P, Q99P et Qmax) pour la période 1991-2020 révèle des tendances significatives à la hausse, avec des pentes de Sen pouvant atteindre +20 $m^3/s/an$ pour Qmax et une rupture marquée en 1993, coïncidant avec une évolution du régime pluviométrique après les sécheresses des années 1980-1990 (Chaibou Begou et al., 2016). L'ensemble de ces constats souligne une intensification des écoulements extrêmes.

Pour pallier les lacunes de mesures in situ et approfondir la compréhension de la dynamique hydrologique, plusieurs modèles de machine learning ont été utilisés afin de prédire le niveau d'eau dans le BRB. Les algorithmes RF, GB, SVM et XGB ont été entraînés à partir des données de la pluie (CHIRPS), de l'évaporation totale, de la profondeur de la nappe et des niveaux d'eau de stations situées en amont (WL_Sof : Niveau d'eau à Sofara, WL_Bou : Niveau d'eau à Bougouni, WL_Dou : Niveau d'eau à Douna et WL_Pan : Niveau d'eau à Pankourou). Une validation croisée en deux étapes (Fold 1 et Fold 2), avec optimisation bayésienne des hyperparamètres (Snoek et al., 2012), a permis de tester la robustesse des modèles. Les résultats montrent que le SVM atteint un NSE supérieur à 0,7 sur le premier volet, tandis que RF obtient un NSE de 0,38 en test puis 0,4 en validation. Lors du second volet, tous les modèles présentent des performances satisfaisantes (NSE entre 0,7 et 0,94), avec, de manière récurrente, une influence majeure de WL_Bou, WL_Sof et Tot_evap. Parallèlement, afin d'évaluer l'évolution future des extrêmes hydroclimatiques, plusieurs modèles CMIP6 (par exemple MPI-ESM1-2-HR, CanESM5, ACCESS-CM2) ont été comparés à CHIRPS via un diagramme de Taylor, révélant que l'ensemble multimodèle (Ensmean) offre souvent le meilleur compromis. Les

scénarios SSP2-4.5 et SSP5-8.5 mettent en lumière des modifications marquées de trois indices de précipitations extrêmes (CDD, CWD, RX1DAY) au cours des prochaines décennies : dans le sud du bassin, les périodes sèches (CDD) pourraient s'allonger sous le scénario SSP5-8.5, tandis que des précipitations plus intenses (RX1DAY) accroîtraient le risque d'inondations rapides. Les ruptures temporelles anticipées varient selon le scénario et l'indice étudié (2033 pour CWD en SSP5-8.5, 2041 pour CWD en SSP2-4.5). Enfin, un indicateur de niveau d'eau maximum (WLmax) destiné à l'évaluation des crues indique une tendance à la baisse, avec des ruptures vers 2036 sous SSP2-4.5 et 2034 sous SSP5-8.5, associées à des pentes de Sen respectives d'environ -0,01 et -0,017 m/an. Les projections suggèrent ainsi une variabilité intra-bassin accrue, alternant entre épisodes de pluies torrentielles et périodes prolongées de déficit hydrique, et mettent en évidence la nécessité d'adapter la gestion de l'eau et l'infrastructure hydrologique pour faire face aux changements climatiques à venir.

Conclusions et Perspectives

Cette recherche a permis de mieux comprendre la variabilité hydroclimatique du BRB et à renforcer les capacités de prévision des fluctuations du niveau d'eau. Elle aborde quatre questions majeures : la précision des estimations satellitaires des précipitations, l'impact des événements hydroclimatiques extrêmes, l'efficacité des modèles d'apprentissage automatique dans la prédiction des niveaux d'eau, et les implications des scénarios climatiques futurs sur les ressources en eau d'ici 2050.

- **Précision des estimations satellitaires de précipitations**

  Les produits satellitaires, notamment CHIRPS, se révèlent fiables pour combler les lacunes des données dans le BRB, offrant une alternative efficace là où les mesures au sol sont limitées par des défis de logistiques ou de sécurité.

- **Impact des événements hydroclimatiques extrêmes**

  Des changements significatifs dans le régime des précipitations et des débits depuis les années 1990 indiquent un risque accru d'inondations et des défis pour l'agriculture et la pêche, soulignant la nécessité d'adaptations dans la gestion des ressources en eau.

- **Performance des modèles d'apprentissage automatique**

  Les modèles tels que Random Forest et Gradient Boosting se sont révélés particulièrement performants pour reproduire les niveaux d’eau historiques, confirmant leur capacité à surmonter les limites mises en évidence par Chaibou Begou et al. (2016) dans le bassin de la Bani, où les modèles hydrologiques classiques, tels que SWAT, peinent à restituer fidèlement la dynamique hydrologique en raison du manque et de l’irrégularité des données.

- **Scénarios futurs et gestion des ressources en eau**

  Les projections basées sur les scénarios SSP2-4.5 et SSP5-8.5 mettent en évidence des effets contrastés sur les ressources en eau du BRB, avec des implications sérieuses pour la gestion future de l'eau, nécessitant des politiques adaptatives et des infrastructures résilientes.

Des recherches ultérieures devraient se concentrer sur l'amélioration des modèles climatiques à une échelle locale plus fine et le renforcement des réseaux d'observation au sol. L'intégration de facteurs socio-économiques pourrait également enrichir les discussions sur l'allocation équitable des ressources en eau, offrant une base pour des politiques de gestion de l'eau plus justes et efficaces. Ces efforts combinés fourniraient aux décideurs des outils nécessaires pour une gestion durable des ressources en eau dans le BRB, favorisant des interventions adaptatives et informées.

List of figures

List of tables

List of content

Chapter 1: General introduction

Water resources in semi-arid regions worldwide are increasingly under pressure due to the combined effects of climate variability, human-induced changes, and socio-political instability (O'Brien et al., 2020). In West Africa, these challenges have particular significance given the dependence of local communities on river systems for agriculture, fisheries, and broader socio-economic activities (Roudier and Mahe, 2010a; Sylla et al., 2018). Among these river systems, the Bani River Basin (BRB) in southeastern Mali stands out for its importance as a tributary of the Niger River, directly supporting agriculture, wetland ecosystems, and fisheries (Mahe et al., 2013; Traore, 2015). However, in-depth scientific attention to the BRB has historically lagged behind that given to larger and better-documented basins such as the Senegal or the main Niger River. Understanding the basin's hydroclimatic trends is essential not only for resource management but also for regional stability, given the area's socio-political vulnerabilities (Kohnert, 2023; Onuoha et al., 2023).

Recent efforts to characterize and predict hydrological fluctuations in the BRB have highlighted several critical gaps in data availability and model accuracy (Anderson and Radić, 2022; Chaibou Begou et al., 2016). Traditional hydrological models, which often rely on consistent ground-based measurements of rainfall and discharge, face substantial challenges due to data scarcity, particularly in conflict-prone or logistically difficult areas (Onuoha et al., 2023). This scarcity is exacerbated by climatic anomalies such as prolonged droughts or unexpectedly intense rainfall events that are becoming more frequent under climate change (O'Brien et al., 2020; Peters et al., 2020). As a result, water managers in regions like the Mopti area of Mali must make critical decisions with incomplete or outdated data, leading to increased vulnerability to both floods and droughts (Ba, 2023).

To address these observational gaps, satellite-derived rainfall datasets such as the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) have emerged as potentially powerful tools (Houngnibo et al., 2023b). These products combine ground station data with remote sensing and reanalysis techniques, offering higher spatial and temporal coverage than would be possible through conventional observation networks alone (Dinku et al., 2010).

However, while data-driven and machine learning (ML) approaches have shown strong potential for improving prediction accuracy under data-scarce conditions, they also present limitations in simulating water-resources management scenarios and capturing the complex feedbacks associated with land-use and anthropogenic changes. In semi-arid and conflict-affected regions, such satellite-based datasets can provide essential information to fill in the missing links of rainfall and river flow data (O'Brien et al., 2020). However, validation of these products is crucial, as local biases may compromise their reliability in complex terrains or areas where ground truth data are limited (Houngnibo et al., 2023c).

Parallel to these advancements in data acquisition, the emergence of machine learning (ML) techniques has opened new avenues for hydrological modeling and forecasting (Hassani et al., 2024, 2020; Tahmasebi and Sahimi, 2021). Models like RF and GB can capture nonlinear relationships within large, heterogeneous datasets, making them particularly suited to the complexities of river basins experiencing variable climatic and anthropogenic pressures (Anderson and Radić, 2022; Xiang et al., 2021). By leveraging satellite-derived rainfall estimates and other remote sensing variables, ML models are increasingly effective in predicting streamflow, water levels, and flood risks in data-scarce regions (Yaseen et al., 2019). Moreover, the flexibility of these techniques allows for the integration of future climate scenarios such as those from the CMIP6 to assess how warming trends may alter rainfall and discharge patterns in the near future (Rameshwaran et al., 2021; Sawadogo et al., 2024).

In the specific context of the BRB, few studies have fully harnessed these combined advantages of satellite data and ML-based methods. While previous research successfully modeled water-level fluctuations using conventional statistical or hydrological models (Chaibou Begou et al., 2016; Mahe, 2009), there remains a strong need to systematically evaluate the performance of satellite rainfall products, analyze long-term hydroclimatic trends, and integrate predictive modeling with future climate scenarios. This integration is particularly urgent in the Mopti region, where socio-political instability complicates ground-based monitoring efforts (Kohnert, 2023; Onuoha et al., 2023).

Against this backdrop, the present thesis aims to strengthen hydrological monitoring and forecasting for the BRB by focusing on four key objectives:

- Evaluating the accuracy of satellite products to mitigate data gaps.

  By comparing satellite-derived rainfall estimates (e.g., CHIRPS) with available ground measurements, this research seeks to quantify the reliability and bias of these datasets in capturing spatio-temporal precipitation variability across the BRB (Dinku et al., 2010; Houngnibo et al., 2023b).
- Analyzing hydroclimatic trends and patterns over recent decades.

  The study will employ robust statistical techniques, such as the modified Mann–Kendall test, to identify trends in precipitation and river flow indices (e.g., RX1DAY, Qmax) over recent decades, thereby offering insights into both gradual and abrupt shifts in the BRB's hydrological regime (Peters et al., 2020).
- Assessing the predictive performance of ML models in replicating historical water levels.

  Advanced ML algorithms, including RF and GB, will be applied to historical climate and hydrological datasets to determine how effectively they can replicate observed water-level fluctuations. This step aims to validate their suitability for operational forecasting in data-constrained environments (Xiang et al., 2021).
- Forecasting future water levels under SSP2-4.5 and SSP5-8.5 climate scenarios.

  Finally, ML-based models will be coupled with CMIP6 climate projections to estimate future water levels, with an emphasis on identifying potential hotspots of hydrological change and informing resource management strategies under different emission pathways (Roudier et al., 2014a; Sawadogo et al., 2024).

By pursuing these objectives, this thesis not only enhances the existing body of knowledge on the BRB but also provides a robust framework for stakeholders including government agencies, local communities, and international organizations to make informed decisions about water resource allocation and climate adaptation. The fusion of satellite-derived data, advanced ML techniques, and scenario-based analyses promises to significantly improve both the accuracy

and the resilience of hydrological forecasts in a region where data scarcity and socio-political uncertainty frequently hamper effective management (O'Brien et al., 2020).

1.1. Context and problem statement

Among the many impacts of climate change, those affecting surface water resources remain particularly concerning due to their direct implications for water availability, agriculture, and disaster management. This challenge is especially pronounced in West Africa, where watersheds are often highly sensitive to both climatic fluctuations and human interventions (Roudier and Mahe, 2010b; Sylla et al., 2015). Despite a growing body of research, uncertainties persist regarding how West African basins respond to changing environmental conditions uncertainties that must be addressed to inform sustainable water management (Coron et al., 2017; O'Brien et al., 2020).

The BRB, a vital tributary of the Niger River, exemplifies these challenges. Previous studies have highlighted its susceptibility to climatic variability and anthropogenic pressures (Chaibou Begou et al., 2016; Mahe et al., 2013), for instance, demonstrated how interannual rainfall variations significantly influence groundwater levels in the BRB. In parallel, (Picouet et al., 2000) linked suspended sediment dynamics to fluctuations in liquid discharges at Banankoro and Douna, while (Roudier and Mahe, 2010b) detailed the basin's declining maximum flows since the 1970s. Collectively, these works underscore how reduced precipitation, hydro-agricultural development, and other human activities adversely affect water resources a vulnerability that is amplified by rising temperatures and erratic rainfall (G. Oyerinde et al., 2017; Sylla et al., 2018).

Further quantifying these pressures, (Paturel et al., 2010) reported substantial decreases in rainfall (about 15–25%) and significant reductions in annual discharges (up to 65%) since the 1970s at Douna, the BRB's outlet. Although physical-based hydrological models such as SWAT have yielded moderate success in capturing these changes (Chaibou Begou et al., 2016), their applications in the BRB have primarily focused on simulating streamflow and evaluating the impacts of land-use and climate variability on water balance components. Chaibou Begou et al. (2016) calibrated and validated SWAT at Douna and Banankoro stations, demonstrating

its potential to reproduce seasonal flow patterns but also revealing significant uncertainty during low-flow periods due to sparse and inconsistent climate and land-cover data.

However, their ability to represent nonlinear climate–water interactions remains limited (Peters et al., 2020). Complicating matters is the ongoing insecurity in the Mopti region, where the Douna hydrometric station is frequently vandalized, undermining the effectiveness of conventional hydrometric monitoring and flood mitigation efforts (Kohnert, 2023; Onuoha et al., 2023). These limitations have led to more frequent and severe flooding in Mopti, amplifying socio-economic vulnerabilities (Ba, 2023).

In response to these challenges, innovative approaches that leverage ML have emerged, offering new pathways for accurately predicting water-level fluctuations under data-sparse and conflict-prone conditions (Xiang et al., 2023). By capturing complex, non-linear interactions among climatic drivers, human interventions, and limited ground-based measurements, ML algorithms can overcome some of the inherent biases in traditional empirical formulas (Anderson and Radić, 2022). Equally important is the use of satellite data, such as CHIRPS, to compensate for observational deficiencies in regions where security risks hinder routine data collection (Houngnibo et al., 2023b). According to (Ahmed et al., 2022a), XGB and Gaussian process regression (GPR) remain underutilized for forecasting river water levels, despite their potential in handling multifaceted hydrological datasets.

Building on this momentum, the present study aims to develop a robust ML-based framework incorporating datasets from satellites, historical records, and socio-economic indicators to accurately forecast water-level fluctuations in the BRB. Such a framework will not only strengthen predictive accuracy but also facilitate adaptive flood management strategies. Ultimately, these insights serve to mitigate the escalating impacts of climate change on water resources and bolster regional resilience, especially in contexts where socio-political unrest and inadequate infrastructural systems pose substantial risks to vulnerable communities.

1.2. State of art

A powerful model for predicting river water-level fluctuations is crucial for effective hazard mitigation, particularly in the context of increasing flood risks (Tikhamarine et al., 2020). In

response, numerous researchers have explored diverse methodologies to address the inherent complexities of water-level variations (Yaseen et al., 2019). However, conventional modeling approaches such as autoregressive or Box-Jenkins models encounter notable limitations. They often require extensive parameter calibration and assume a substantial reservoir of historical data, which is not always available or sufficiently detailed (Kumar et al., 2021). Physics-based models add another layer of complexity by requiring multiple assumptions and a comprehensive understanding of the system's dynamics (Saldanha Do Carmo et al., 2020). For instance, developing a purely physical model to evaluate river floodplains can be costly and data-intensive (Teng et al., 2023).

Over time, numerical approaches have emerged as a way to alleviate some constraints of physics-based models (Wu et al., 2014). Nonetheless, errors in numerical modeling persist, often stemming from incomplete representations of the physical processes involved (Guan et al., 2020). Recent advancements in data-driven techniques have shown promise in addressing these issues, as demonstrated by the application of various machine learning (ML) algorithms for different hydrological parameters. For example, (Allawi et al., 2018) employed radial basis functions and support vector machines to predict reservoir evaporation across daily, weekly, and monthly time horizons, achieving high levels of accuracy. Similarly, (Kamel et al., 2021) used ML methods to model subsurface evaporation, and (Hipni et al., 2013) found that support vector machines outperformed ANFIS for daily water-level prediction in the Klang Gate reservoir.

Despite their generally higher accuracy compared to classical techniques, ML models require proper hyperparameter tuning (e.g., weights, activation and transfer functions) to achieve optimal performance (Ahmed et al., 2022a). Recognizing this need, numerous studies have combined ML algorithms with meta-heuristic optimizers, such as genetic algorithms or particle swarm optimization, to enhance river-discharge predictions (Banadkooki et al., 2019; Osman et al., 2021; Pham et al., 2021; Rezaei et al., 2020; Tikhamarine et al., 2020). While these hybrid approaches generally improve accuracy, some optimization techniques suffer from low or unstable convergence rates and may increase computational demands (Shahabi et al., 2020).

Tree-based regression models have recently gained attention as potentially less complex yet effective alternatives to hybrid ML approaches (Ghazvinian et al., 2019). (Ridwan et al., 2021) showed that boosted decision trees and decision forests could surpass classical models in daily rainfall prediction. In water-level applications, boosted linear and decision tree regressions outperformed traditional neural networks and radial basis function methods (Sapitang et al., 2020). XGB, in particular, has demonstrated both high accuracy and computational efficiency in modeling groundwater fluctuations (Osman et al., 2021). As an ensemble method anchored in gradient boosting, XGB often excels at capturing sudden changes in hydrological parameters. Machine learning techniques have also been extended to broader climate-related challenges, such as flood and storm prediction, particularly in coastal areas increasingly exposed to socio-environmental complexities (Bamal et al., 2024). In a similar vein, (Xu et al., 2020) employed ML-based downscaling to improve the accuracy of regional climate simulations for the upper Han River basin, highlighting the value of data-driven methods at the watershed scale. Furthermore, (Adounkpe et al., 2020) achieved Nash–Sutcliffe efficiencies exceeding 0.93 by applying deep learning models to streamflow prediction in Niamey, underlining the effectiveness of neural networks in hydrological applications. These findings collectively emphasize the growing significance of machine learning and specifically ensemble or optimized ML algorithms as flexible, powerful tools for addressing the complexities and uncertainties of water-level prediction in a changing climate.

1.3. Research questions

In view of the above-mentioned problems, it is imperative to find the answer to the following questions. To do so, we will use during this research work recorded data from hydrometeorological services, agricultural services, climate scenarios in order to answer the following questions :

- How do satellite rainfall estimates compare to ground observations in terms of accuracy and reliability?
- How have changes in extreme hydroclimatic events affected the BRB's hydrology?

- Which machine learning models most effectively reproduce water level fluctuations under data-scarce and hydroclimatically variable conditions in the Bani River Basin (BRB)?
- What do the climate scenarios tell us about the evolution of water levels in the BRB by 2050?

1.4. Thesis aim and objectives

To develop a robust and comprehensive predictive framework using advanced ML algorithms to forecast water level fluctuations in the BRB under changing climatic conditions. This aim is grounded in addressing critical gaps in current hydrological modeling approaches, particularly in regions like the BRB where data scarcity, climate variability, and anthropogenic influences compound water management challenges. By leveraging machine learning techniques, the thesis seeks to enhance predictive accuracy, mitigate flood risks, and establish a generalizable methodology that can be applied to similar river basins affected by climate change.

The specific objectives of this study are outlined below, detailing the key steps and focus areas undertaken to achieve the overall aim of the research:

- Evaluating the accuracy of satellite products to mitigate data gaps;
- Analyzing hydroclimatic trends and patterns over recent decades;
- Assessing the predictive performance of ML models in replicating historical water levels;
- Forecasting future water levels under SSP2-4.5 and SSP5-8.5 climate scenarios.

1.5. Novelty

This study introduces a novel hybrid framework that merges satellite-derived datasets and advanced machine learning algorithms for water-level forecasting in semi-arid, conflict-prone environments like the BRB. By integrating satellite based rainfall data, ML algorithms, it circumvents the constraints of scarce ground-based observations. Furthermore, we incorporate local socio-political dimensions, acknowledging that instability can disrupt conventional data collection. This integration ensures model adaptability and continuous forecasting even under

data-limited situations. Additionally, by applying CMIP6 climate scenarios, the research does not merely improve near-term predictions but also provides decision-makers with long-range tools for adaptation. Such scenario-based modeling highlights critical breakpoints and potential extremes. Lastly, the framework's transferability stands out: other semi-arid regions facing similar climatic, infrastructural, or security hurdles can replicate these methods. Hence, the synergy of satellite data, robust computational techniques, and scenario-based forecasting offers a groundbreaking pathway for sustainable water management under compound uncertainties. This integrated perspective addresses multiple layers of uncertainty, bridging data gaps while providing proactive guidance for local stakeholders and policymakers.

1.6. Scope of the thesis

This research focuses on improving water-level forecasting and management within the BRB, a region significantly affected by climatic variability and socio-political uncertainties. By integrating satellite-based rainfall products (e.g., CHIRPS), traditional hydroclimatic observations, and ML algorithms, the study addresses the pressing need for accurate, data-driven decision-making tools. The thesis encompasses several dimensions: evaluating the reliability of satellite-derived datasets, applying advanced ML models to predict water level fluctuations, and analyzing the implications of future emission scenarios on water availability. Through close collaboration with local hydrological authorities in Mopti, the research aims to deliver actionable insights that can enhance resource allocation, guide infrastructural investments, and improve resilience in the face of climate extremes and socio-political challenges.

1.7. Expected results and benefits

This study integrates satellite-derived rainfall data with advanced machine learning models to enhance water-level forecasting in the BRB. By doing so, it will fill critical data gaps and enable more robust water management strategies under complex climatic and socio-political conditions. Specifically, the research is expected to:

- **Improve rainfall data quality**

Validate and refine satellite-based rainfall products (e.g., CHIRPS) to compensate for sparse ground measurements, thereby providing more reliable inputs for hydrological modeling.

- **Enhance water-level forecast accuracy**

Employ ML algorithms such as RF, GB, XGB and SVM to capture nonlinear hydrological dynamics, ultimately achieving more accurate and timely river-level predictions.

- **Offer insights on hydroclimatic trends and extremes**

Identify key shifts in rainfall patterns and extreme flow events, informing adaptive measures for both flood mitigation and agricultural water planning.

- **Strengthen adaptive water management**

Generate evidence-based recommendations that can be directly applied by local authorities and stakeholders to reduce climate-related risks and secure livelihoods in the region.

- **Provide scenario-based future projections**

Forecast potential changes in water availability under different climate pathways (e.g., SSP2-4.5, SSP5-8.5), guiding long-term infrastructure development and ecosystem conservation efforts.

- **Establish a transferable modeling framework**

Demonstrate a scalable approach that can be replicated in other semi-arid and conflict-prone regions, promoting broader adoption of innovative data-driven methods in water resource management.

Chapter 2: Study area

In this chapter, we delve into a comprehensive examination of the BRB, a pivotal hydrological and ecological system within West Africa. This analysis encompasses several critical dimensions necessary to understand the basin's physical and socio-economic attributes thoroughly. We initiate with a geographical positioning (section 2.1) of the BRB, delineating its coordinates, the countries it traverses, the main cities it impacts, and its overall catchment area. Further, we explore the basin's boundaries and outline its expansive reach across Mali, Côte d'Ivoire, and Burkina Faso.

Following the geographical overview, we detail the basin's relief (section 2.2), discussing the topography including the varied elevations from mountains to plains, and the topographic features that influence the hydrological characteristics of the river. This segment also addresses the average altitude and key topographical variances that affect water flow and availability within the basin.

Vegetation within the basin is explored next (section 2.3), where we classify the types of vegetation present, ranging from lush savannas to arid desert patches. This section elaborates on the dominant flora species, their distribution, and the impacts of anthropogenic activities on these natural habitats.

The climatic conditions of the BRB are meticulously analyzed (section 2.4), covering aspects such as annual precipitation patterns, temperature averages, and the distinct seasons. We also delve into the observed climatic fluctuations over recent decades and assess the implications of climate change on this sensitive region.

In the hydrographic overview (section 2.5), we concentrate on the intricate network of rivers and tributaries within the basin, emphasizing the BRB's role and its significant tributaries. Discussions here include aspects of river discharge, the regime of floods and droughts, and the infrastructure like dams and reservoirs that are crucial for effective water management.

Soil types and land usage in the basin are discussed (section 2.6), highlighting the primary soil groups and how land is utilized, from agricultural practices to conservation areas. This helps in understanding how soil and land management strategies are being implemented to support

sustainable development. We then present demographic characteristics (section 2.7) of the populations residing within the basin, including population density, geographical distribution, and major societal trends. This section also touches upon migration patterns that affect the socio-economic landscape of the region. Lastly, the section on environmental, social, and economic activities (section 2.8) discusses the main economic activities such as agriculture, livestock, and energy production, and their interplay with the environmental dynamics of the basin. we synthesize the key elements discussed in this chapter, emphasizing the complex interdependencies between the physical geography, climate variability, and human factors influencing the BRB. We also highlight the ongoing challenges and potential opportunities for fostering sustainable development within this vital water catchment area.

2.1. Localization

This study focuses on the BRB (Figure 2), which flows into its outlet in Mopti, central Mali. The river basin spans approximately 129400 km², covering parts of Mali, Côte d'Ivoire, and Burkina Faso, with coordinates between 9°10'13"N to 14°05'44"N latitude and 8°05'16"W to 3°46'17"W longitude. While the majority of the basin lies within Mali, parts of it reach into Côte d'Ivoire and Burkina Faso (Chaibou Begou et al., 2016). The Bani River is formed by the Baoulé and Bagoé rivers, which originate in Ivory Coast. It then receives its main tributary, the Banifing from Kouoro on the right bank before reaching the Douna station and entering the inner delta. In the delta, the Bani is joined by the Banifing from San and then the Yamé from Bandiagara, flowing to Mopti where it merges into the Niger River (Paturel et al., 2010).

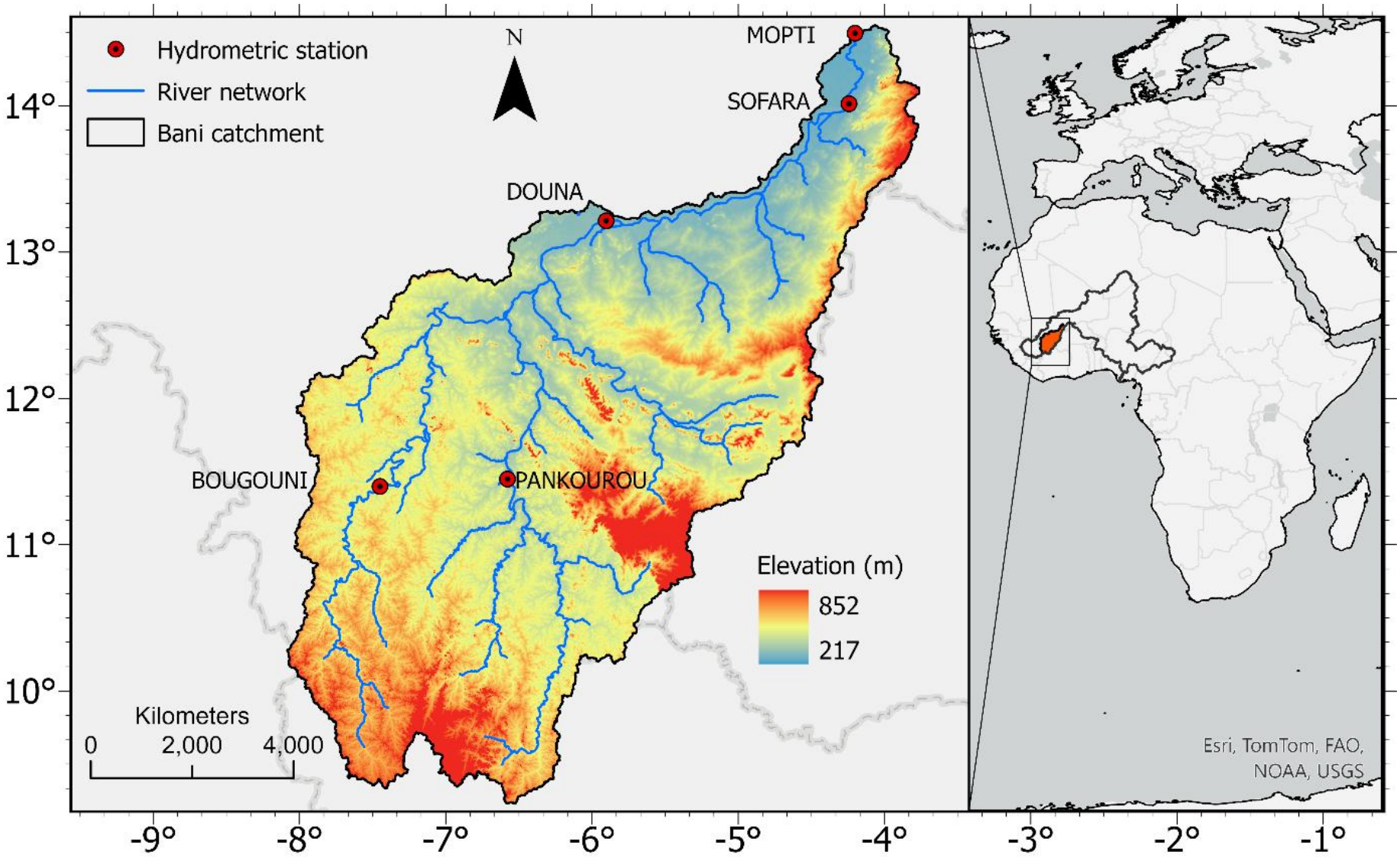

Figure 2: Bani catchment in western Africa shared by Mali, Côte d'Ivoire, and Burkina Faso.

2.2. Relief

The BRB is an integral part of the Upper Niger Basin, marked by its distinct geographical boundaries and varied topography. It is bordered to the east by the meridian near San and to the west by the meridian near Bamako. The southern limit slightly overlaps into the territories of Côte d'Ivoire and Burkina Faso, while the northern boundary runs roughly parallel to the Niger River, positioned about forty kilometers to the south of it. The topography of the BRB is characterized by its generally subdued relief, with more pronounced elevations found particularly along the southern edge of the basin, in the high Bagoé massif west of Boundiali. Here, the peaks reach up to approximately 852 meters (Figure 2). The essential topographic feature of the basin is a very gentle slope descending northeastward from elevations of about 400 meters to 300 meters. The only significant variations in this gradient are found in the southern parts of the basin, where latéritic buttes with flat tops are heavily dissected by the hydrographic network. To the north, along the Bougouni-Sikasso line, the landscape features flat-topped sandstone plateaus.

This elevation map vividly illustrates these features, highlighting the basin's diverse terrain with a color gradient that accentuates areas of higher elevation in red and orange tones, which transition to cooler blues as the elevation decreases. The network of rivers is clearly delineated, underscoring their role in shaping the basin's landscape and influencing its ecological and hydrological dynamics. Hydrometric stations positioned at key locations within the basin provide essential data for monitoring the environmental and hydrological changes over time.

2.3. Vegetation

The BRB lies entirely within a savanna region. In this ecosystem, herbaceous vegetation dominates: it grows rapidly during the rainy season, then dries out or dies off by the end of its growth cycle. As the dry season lengthens and becomes more intense, trees and shrubs become progressively rarer, though they still appear alongside the grasses. In the southern portion of the basin particularly in the area extending into Côte d'Ivoire the landscape takes the form of a classic Guinean woodland savanna. Here, shrubs form a fairly dense layer, and the grassy

understory is extensive. Also small patches of gallery forest and forest islets, though these are not widespread. This type of savanna stretches northeastward along the basin, eventually reaching the Sikasso region. It safeguards the soil by slowing runoff, thereby increasing the land's water retention capacity (Vreysen et al., 2013).

Progressing northward, the savanna thins out; gallery forests disappear, and vegetation becomes more xerophytic (adapted to dry conditions). In this more scattered form, it offers only limited resistance to runoff especially early in the rainy season before the grasses regenerate.

In closing this overview of the Bani watershed's physical environment, it is worth noting that the basin's morphology integrates several elements: relatively gentle topography, substantial flood-water losses, and plant cover that from August onward effectively mitigates runoff while curbing erosion. Throughout the region, the slopes are gentle and uniform (Kouyaté et al., 2025).

Together, these factors exert a moderating effect on runoff. The sizable water losses postpone the high-flow period by about one to one and a half months after the usual late-August peak rainfall. Meanwhile, once floodwaters recede, the floodplains release only part of what they had absorbed.

2.4. Climat

The study area is influenced by tropical climate dynamics, experiences a stark seasonal variation driven primarily by the interactions of different air masses and the movement of the Intertropical Convergence Zone (Balcázar et al., 2022). This zone, a critical climatic boundary, shifts its position throughout the year, impacting the precipitation patterns and climate conditions across the basin. During the boreal summer, the Intertropical Convergence Zone (ITCZ) reaches its northernmost position, bringing moist, unstable maritime air from the south, which contrasts with the dry, stable continental air originating from the Saharan anticyclone. This setup results in a distinct wet season characterized by heavy rains and a dry season marked by reduced precipitation.

Further insights into the regional climate are provided by examining specific climatic factors such as wind patterns, temperatures, and relative humidity across various meteorological

stations including Bougouni, Sikasso, Koutiala, Ségou, San, and Mopti. The wind regime, particularly dominated by the Harmattan from the northeast and the moist southwesterly monsoons, illustrates the dynamic climatic interactions at play (Ibrahim et al., 2017). Temperature data indicates a general increase from south to north across the stations, with seasonal highs peaking just before the onset of the rains and lows during the dry season. Relative humidity follows a similar seasonal pattern, with the highest values recorded during the peak of the rainy season in August and the lowest during the dry season, reflecting the broader climatic influences affecting the basin. This nuanced understanding of the climatic conditions in the BRB highlights the complex interplay of atmospheric dynamics that govern the seasonal and spatial variations in weather patterns, essential for managing the region's agricultural and water resources (Paturel, 2014).

2.5. Hydrography

The Bani River is formed by the confluence of the Baoulé and Bagoé rivers, which originate in the hills dividing the Niger Basin from the coastal rivers of Côte d'Ivoire. This drainage network covers southern Mali and northern Côte d'Ivoire, as illustrated in figure 3.

The Baoulé River sources near Odienné at an elevation of 450 meters, gently flowing northward towards the Niger River. From its origins, the river is characterized by meanders and floodplains, which become evident upstream of Odienné. After traveling approximately 30 kilometers, the Baoulé shifts into a south-north trajectory with a slope of 40 centimeters per kilometer, where it meets its first tributary, the Banifing. The slope then reduces significantly to 20 centimeters per kilometer, enhancing the meandering as it proceeds. Further downstream, the Baoulé picks up another tributary, the Dégou, on its right bank. As it passes through Bougouni, 500 kilometers from its source, the floodplain widens dramatically to between 800 and 1500 meters. The river continues to collect tributaries, including two more Banifings, one on the right and another on the left bank. After navigating through a more confined course, the Baoulé turns west-east and meets the Bagoé just past Doïla, 800 kilometers from its origin.

The Bagoé River begins south of Boundiali at 600 meters above sea level, descending through a plain in a south-north direction, quickly developing into a river with broad floodplains and a

winding course. Midway, it collects the Bafini on the right bank and the Kankelaba and another Banifing on the left bank. The gradient lessens and the meanders become more pronounced as the Bagoé stretches over 700 kilometers to join the Baoulé.

Figure 3: Drainage network of the BRB

2.6. Soil and land use

According to the Food and Agriculture Organization (2003) (Ruelland et al., 2012), the primary soil groups are predominantly made up of Luvisol, Acrisol, and Nitosol (Figure 4a). As per the USGS Global Land Cover Characterization (GLCC) version 2.0 (Loveland et al., 2000), the predominant form of land use is agricultural land, with savannah and forest following in extent (Figure 4b). The figure 4a showcases the diverse soil types found throughout the Bani catchment. Each soil type is characterized by unique properties that influence their use and

management. Ferric Acrisols, which are widespread across the catchment, are deeply weathered and rich in iron and aluminum oxides, posing challenges for agricultural productivity without significant amendment and careful management. Orthic Acrisols and Plinthic Acrisols, which appear in various patches, have better drainage compared to Ferric Acrisols and support a broader range of agricultural activities, although they still require careful management to prevent nutrient depletion. Eutric Cambisols and Gleysols, more fertile and with better water retention capabilities, are ideal for more intensive agricultural practices. Lithosols, scattered throughout the region and characterized by shallow depth and stony texture, are found mainly on steeper slopes and are generally less suitable for cultivation but may support certain types of pasture or forest. Luvisols and Nitisols are deeper and well-drained soils that are among the most fertile in the region, highly suitable for various agricultural uses due to their higher nutrient content and good structure. The land use map (Figure 4b) divides the Bani catchment into three primary categories that reflect the dominant forms of land utilization. Agricultural land dominates the landscape, depicted in yellow, underscoring the region's dependence on agriculture, which varies in intensity and method depending on the underlying soil type and water availability. Forest areas, shown in green, are primarily located in zones where the soil and microclimatic conditions support such vegetation. These areas are essential for maintaining ecological balance, influencing soil and water conservation, and providing habitat for diverse wildlife. Savannah, indicated in light green, is typical for areas with less intensive land use, often used for grazing and integral to the ecological diversity of the catchment, supporting a mix of natural vegetation and wildlife.

The interaction between soil types and land use patterns is evident from the maps. For example, agricultural lands on Ferric Acrisols may require different management strategies, such as the addition of organic matter and regular soil testing, to maintain soil health and productivity compared to those on Eutric Cambisols. Similarly, the presence of forests and savannahs in areas with specific soil types can suggest regions where natural vegetation helps prevent erosion and maintain ecological stability.

This comprehensive examination of the Bani catchment's soil and land use maps helps stakeholders understand the complexities of land and soil resource management. Such understanding is crucial for developing sustainable agricultural practices, conserving natural resources, and planning future land use in a way that balances productivity with environmental conservation.

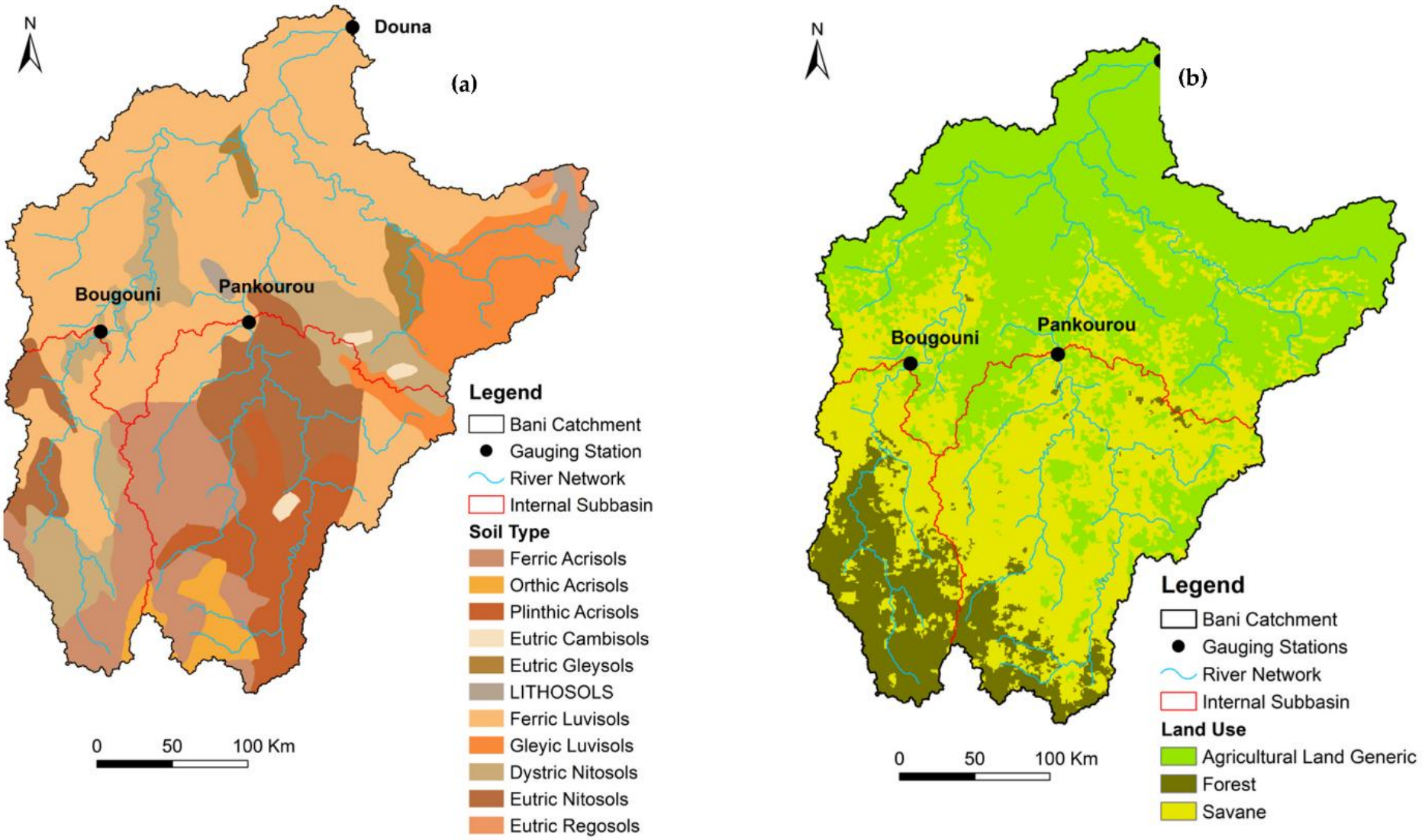

Figure 4: (a) Soil attributes and (b) land use categories of the Bani catchment (Chaibou Begou et al., 2016)

2.7. Demography, environmental, social and economic activities

- Population

The Figure 5 illustrating the Bani catchment displays population distribution based on data from the WorldPop Global Project, which estimates residential populations within precise 100x100m grid squares from 2000 to 2021. This high-resolution data, derived through a RF-based dasymetric redistribution method, is informed by recent census figures and various geospatial layers, ensuring accurate and detailed demographic mapping. The total population of the Bani catchment is estimated at 5,957,583, with the map's visual gradations highlighting areas of varying population density. The WorldPop dataset, peer-reviewed and constructed with transparent methodologies, is crucial for multiple applications, including infrastructure planning, environmental management, and emergency response (Linard et al., 2012). The availability of this data on an open-access platform supports a wide array of research and policy-making initiatives, enhancing efforts in sustainable development and resource allocation. Funded primarily by the Bill and Melinda Gates Foundation and developed in collaboration with several academic institutions, this dataset represents an invaluable resource for understanding and managing population dynamics effectively.

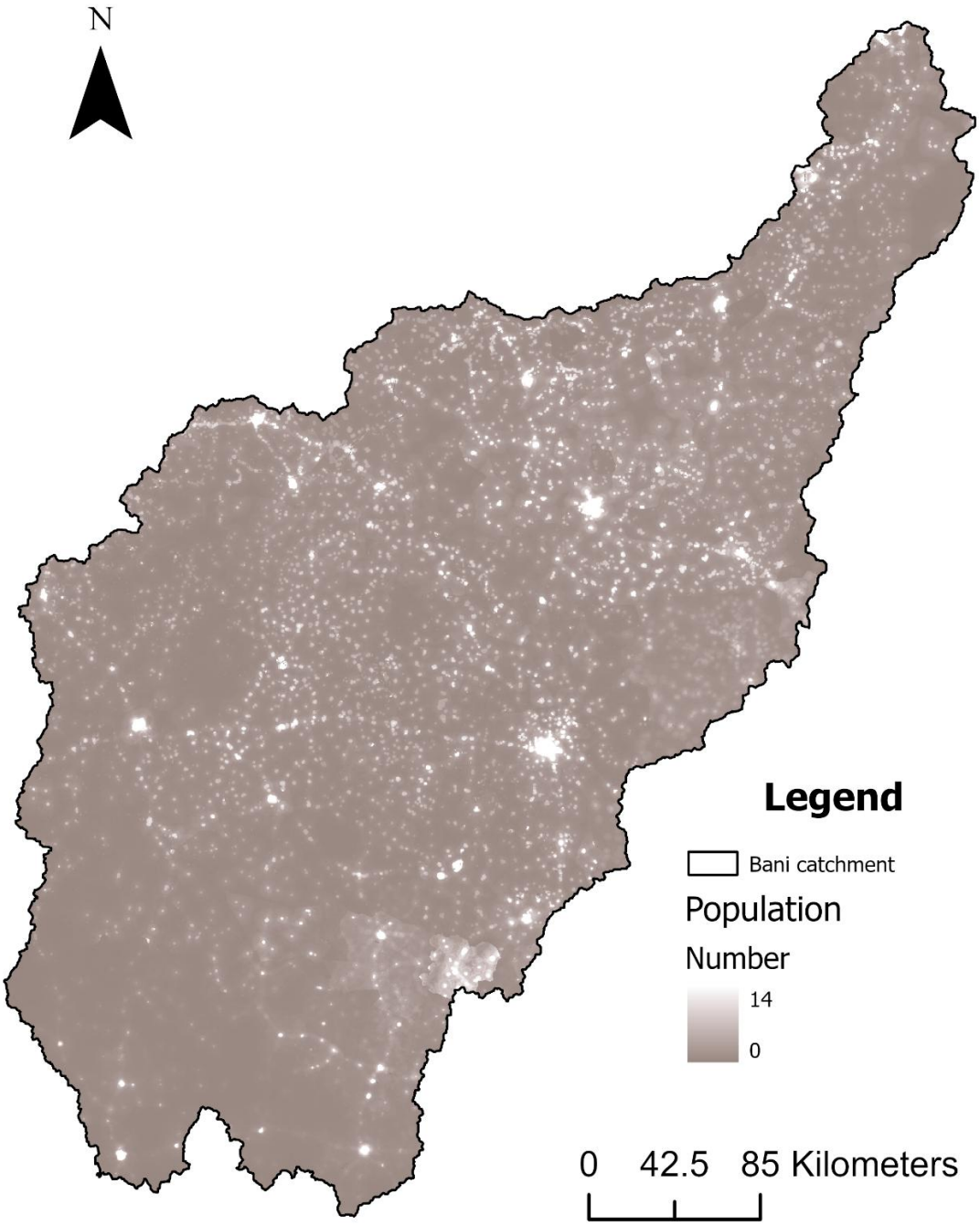

Figure 5: Population distribution in the Bani catchment Based on WorldPop Data.

- Social and economic activities

The BRB serves as a crucial tributary of the Niger River, plays a vital role in the environmental, social, and economic fabric of the region. The basin's water availability is highly variable, often limited during dry seasons, which poses significant risks to agricultural productivity and food

security. This variability is exacerbated by climate change, which intensifies the seasonal fluctuations of river flow and complicates water management efforts.

In terms of infrastructure, the region faces a pressing need for development, particularly in expanding irrigation schemes and constructing dams and reservoirs (Roudier et al., 2014a). Such projects are essential to ensure consistent water availability, supporting both agricultural activities and the daily needs of the population throughout the year.

Energy access within the basin is notably limited, with only 26% of the population having access to electricity(Toure et al., 2016). This scarcity is more acute in rural areas, affecting approximately 11 million people. The primary sources of power generation in the area include hydropower from the Sélingué and Manantali dams, which contribute 90% of the available electricity (Touré et al., 2016). Additionally, there is a heavy reliance on biomass firewood, charcoal, and agricultural residues which underscores the need for more sustainable energy solutions.

The agricultural sector, deeply intertwined with the hydrological and energy challenges, is crucial for the local economy. Seasonal and climatic changes impact not only water availability but also the livelihoods of those who depend on agriculture and livestock. The ongoing population growth in Mali, particularly the high percentage of youth, adds pressure to improve water and energy infrastructures to enhance people's quality of life and economic stability.

Stakeholders from various sectors, including local government agencies, international donors like the World Bank and the European Union, and numerous local committees, are actively engaged in addressing these challenges. Projects funded by international grants, aim to develop and implement solutions that address these multifaceted socio-economic issues, focusing on sustainability and long-term resilience for the BRB.

2.8. Partial conclusion

This chapter has provided a comprehensive exploration of the BRB, a region pivotal not only for its hydrological and ecological significance but also for its profound socio-economic impact across West Africa. Through a detailed examination that spans geographical delineation, topographical features, vegetation types, climatic conditions, hydrographic characteristics, soil

and land use patterns, and demographic insights, we have painted a vivid picture of a basin that is as diverse in its challenges as it is in its opportunities.

From the varying altitudes affecting hydrological flows to the intricate river networks that sustain a wide array of biodiversity, the BRB stands as a testament to the dynamic interplay between nature and human activity. The pressures of climate change, population growth, and socio-economic developments call for a concerted effort to manage the basin's resources sustainably. Effective water management strategies, informed by the basin's comprehensive hydrographic and climatic analysis, are crucial for mitigating the impacts of variability in water availability and the intensifying extremes of droughts and floods.

The demographic trends and the patterns of land use underscore the necessity for integrated approaches to regional planning that consider both the environmental sustainability and the socio-economic well-being of the basin's inhabitants. The ongoing initiatives aimed at improving irrigation, enhancing energy access, and fostering economic stability are steps toward addressing the intricate mosaic of environmental, social, and economic challenges.

As we conclude this chapter, it is clear that the BRB is a region where every aspect of its physical geography informs and influences the lives of its people. The insights garnered here should guide future policies and research, aiming not only to preserve the ecological integrity of the basin but also to enhance the quality of life for all its residents. Through adaptive management and informed decision-making, the BRB can continue to be a source of life and a cradle of human civilization in West Africa.

Chapter 3: Materials and Methods

This chapter outlines the methodological framework adopted to fulfill the specific objectives of this research, highlighting both data-centric and analytical approaches. It begins with Section 3.1, where historical rainfall data from satellite-based (PERSIANN-CDR and CHIRPS) and in-situ sources are collected, preprocessed, and integrated for preliminary analyses. A reliable Reference Rainfall Product (RRP) is then developed and validated through geostatistical techniques, facilitating a robust evaluation of satellite-derived estimates.

In Section 3.2, the focus shifts to identifying and characterizing extreme hydroclimatic events, including both rainfall and river discharge indices, which form the cornerstone for understanding flood and drought risks. This section also details the trend and change-point detection methods (Modified Mann–Kendall, Pettitt's test, and SNHT) employed to discern long-term patterns and abrupt shifts in the basin's hydrological regime.

Subsequently, Section 3.3 introduces the implementation of four machine learning algorithms XGB, SVM, RF, and GB for predicting water level fluctuations based on a carefully configured dataset incorporating precipitation, evaporation, water-table depth, and upstream water levels. A cross-validation scheme ensures that the predictive models are tested on multiple temporal segments to enhance their generalizability and reliability.

The section 3.4 discusses future hydroclimatic projections derived from different CMIP6 climate scenarios (SSP2-4.5 and SSP5-8.5). Rainfall and temperature outputs from a suite of global models are analyzed to forecast changes in key indicators such as consecutive dry/wet days, maximum daily rainfall, and peak water levels. This integrated approach, combining historical data, state-of-the-art interpolation, machine learning, and climate scenario modeling, provides a comprehensive understanding of the basin's past, present, and potential future hydroclimatic conditions.

3.1. Data collection and preliminary analysis for the first specific objective

Historical rainfall data were collected from two primary sources, PERSIANN-CDR and CHIRPS, covering the period from 1983 to 2000. PERSIANN-CDR provides daily rainfall

estimates at a 0.25° spatial resolution, while CHIRPS combines satellite imagery with in-situ station data to improve accuracy. Additionally, observed rainfall data from meteorological stations within the Basin Region were included for validation. These datasets were preprocessed to ensure temporal alignment and consistency across the different sources. The combination of these datasets forms the basis for evaluating rainfall variability and hydrological modeling within the study area.

- Construction of RRP

To develop a reliable reference rainfall product, we employed a geostatistical approach using data from 26 observed rain gauge stations distributed across the study area. The methodology began with variogram analysis, a key step in understanding the spatial continuity of rainfall patterns (Jeong et al., 2020). Variograms help to explore how rainfall values change over distance, allowing us to assess whether the data are spatially autocorrelated, which is crucial for accurate spatial interpolation. The semi-variogram, which plots the variance of rainfall differences as a function of distance between stations, was used to identify spatial trends and potential anomalies in the data (Oliver and Webster, 2014). One common issue we encountered during this analysis was the "nugget effect", which arises due to measurement errors or spatial variations at a scale smaller than the distance between stations (Jing et al., 2014). Following this, we proceeded to model fitting where various theoretical models such as spherical, exponential, and Gaussian models were tested to find the best fit for our empirical variogram. These models are essential for predicting rainfall at unsampled locations. Once a suitable model was chosen, we performed cross-validation. This involved the "Leave-One-Out" method, where we removed one observation at a time and used the remaining data to predict its value. We then compared the predicted value with the actual observation to calculate error statistics. This process helped us to refine our model and ensure the accuracy of the interpolation. The final step involved applying ordinary Kriging, a geostatistical technique that uses the spatial relationships described by the variogram to estimate rainfall values at unsampled grid points across the study area (Rezaei et al., 2020). The result was a spatially continuous rainfall dataset,

our reference rainfall product that is essential for comparing satellite rainfall products and enhancing our understanding of regional rainfall variability.

- Evaluation of satellite rainfall products

In this investigation, we used two precipitation estimation products: the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR). CHIRPS is a global precipitation estimation product that integrates satellite data and ground observations (Table 1), unlike PERSIANN-CDR, which is exclusively satellite-based product.

- CHIRPS

CHIRPS is a satellite-based precipitation estimation product designed for global coverage. The CHIRPS algorithm integrates infrared-based precipitation data, calibrated using Tropical Rainfall Measuring Mission (TRMM) data, and reanalysis model outputs (Funk et al., 2015). By merging satellite observations with ground-based station data, CHIRPS produces a gridded dataset with a spatial resolution of 0.25°. This dataset spans from 1981 to the present, offering monthly time series for applications such as drought monitoring and water resource management. The CHIRPS process occurs in two phases, incorporating both historical and real-time infrared precipitation estimates to provide comprehensive and continuous precipitation records essential for long-term hydrological assessments (Maidment et al., 2017).

- PERSIANN-CDR

PERSIANN-CDR was developed by the Center for Hydrometeorology and Remote Sensing at the University of California, Irvine. The PERSIANN algorithm employs artificial neural network techniques to calibrate infrared (IR) satellite images with microwave (MW) data for more accurate precipitation estimates. Initially, the algorithm relied solely on IR data from geostationary satellites, but it now incorporates MW data from instruments such as the TRMM Microwave Imager (Tapiador et al., 2004). The IR data used in PERSIANN-CDR are sourced from geostationary satellites, including GOES-8, GOES-9/10, GOES-12, GMS-5,

METEOSAT-6, METEOSAT-7, and VIRS. PERSIANN-CDR is a post-processed product based on archived IR measurements, with its primary input derived from the GridSat-B1 dataset, part of the International Satellite Cloud Climatology Project. This dataset covers the period from 1980 to 2019, combining data from multiple geostationary satellites. Additional inputs include monthly precipitation data from the Global Precipitation Climatology Project, allowing PERSIANN-CDR to produce long-term, high-resolution precipitation records that are valuable for hydrological and climatological studies (Sadeghi et al., 2019).

Table 1: Main characteristics of the satellite products (IRI/LDEO Datathèque du Climat).

Name	Temporal Resolution	Spatial Resolution	Coverage Area	Input Data
PERSIANN-CDR	24 h	0.25°	60 N - 60 S	GridSatB1 + IRWIN
CHIRPS	24 h	0.25°	50 N - 50 S	IR + Ground Observations

- Data validation and quality control

To validate the satellite-derived rainfall data, statistical indicators such as Mean Absolute Error (MAE), Normalized Root Mean Square Error (NRMSE), and Percent Bias (PBIAS) were employed. These metrics quantify the agreement between observed and satellite-estimated rainfall, ensuring the reliability of the CHIRPS and PERSIANN-CDR datasets. The validation process ensured that the satellite data could be confidently used for further hydrological analyses. This approach highlights the importance of integrating multiple data sources to enhance the accuracy of rainfall estimation.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2} \quad (1)$$

RMSE quantifies the square root of the average of the squares of the errors. The error here is the difference between the observed values (O_i) and the predicted values (P_i).

$$\mathrm{NRMSE} = \frac{RMSE}{\max(O) - \min(O)} \quad (2)$$

NRMSE is the RMSE normalized by the range of observed data, which helps to scale the error to the range of the data. This normalization makes it easier to compare the RMSE across datasets with different scales or units.

$$\text{PBIAS} = \left(\frac{\sum_{i=1}^{n}(O_i - P_i)}{\sum_{i=1}^{n} O_i}\right) \times 100 \quad (3)$$

PBIAS measures the average tendency of the predicted values to be larger or smaller than their observed counterparts. It is expressed as a percentage. A positive PBIAS indicates an underestimation by the model, while a negative PBIAS indicates an overestimation. These statistical indicators collectively provide a robust evaluation of the agreement between observed and satellite-derived rainfall data, ensuring their suitability for hydrological and climate-related analyses.

3.2. Analysis of extreme rainfall and river flow indices for the second specific objective

This study analyzes 30 years of data (1991–2020) for the BRB. Precipitation data were sourced from the CHIRPS dataset known for its high resolution (0.05°) and validated reliability for West Africa (Ndiaye et al., 2023). CHIRPS effectively captures the rainfall variability across diverse climatic zones, making it suitable for hydrological modeling. The observed rainfall from 2011 to 2019 was collected from Agence Nationale de la Météorologie du Mali. Discharge data were obtained from the Mopti station, the basin's outlet, provided by the Direction Nationale de l'Hydraulique du Mali. This station's records reflect the natural flow regime, as the basin remains free from significant hydraulic infrastructure. These datasets enable robust analysis of the hydrological dynamics in the study area.

- Methods

- Assessment of observed and CHIRPS rainfall data

Between 2011 and 2019, the quality of the CHIRPS data was assessed by comparing it with rainfall measurements from four stations located along the BRB. Monthly variations in the

satellite estimates were examined alongside the station records. The statistical indicator employed to evaluate the accuracy of CHIRPS was the Nash–Sutcliffe efficiency (NSE).

$$\mathrm{NSE} = 1 - \frac{\sum(\mathrm{Pr}_1 - \mathrm{Pr}_2)^2}{\sum\left(\mathrm{Pr}_1 - \overline{\mathrm{Pr}_2}\right)^2} \quad (4)$$

P_{r1} corresponds to the station-measured rainfall, P_{r2} corresponds to the satellite-derived rainfall, $\overline{\mathrm{Pr}_2}$ corresponds to the mean of satellite-derived rainfall. The NSE ranges from negative infinity to 1; an NSE of 1 indicates perfect agreement between the satellite estimates and station data, while an NSE of 0 suggests that using CHIRPS yields about the same performance as simply taking the average observed value.

- Overview of selected extreme rainfall and river flow indices

The rainfall and river discharge extremes were examined from 1991 to 2021 using specific indices to capture their variability and intensity. The rainfall indices included the extremely wet days (R99P), very wet days (R95P), simple daily intensity (SDII), maximum 1-day rainfall (RX1DAY) and maximum 5-day rainfall (RX5DAY), as shown in Table 2. For the river discharge, three indices were used to describe extreme flow conditions: peak flow (Qmax), high-flow days (Q95P), and very-high-flow days (Q99P). These metrics, calculated over the same period, are crucial for understanding hydrological dynamics during floods and droughts (De Luca et al., 2020; Ndiaye et al., 2023). Additionally, the standardized flow index (SFI) was computed to assess the annual flow variations, identifying years of surplus or deficit. The standardized flow index is defined mathematically as:

$$I_i = \frac{Q_i - \overline{Q}}{\sigma} \quad (5)$$

Table 2: Classification of drought and wet conditions using the standardized flow index.

Values	**Class**
$I_i \geq 2$	Extremely wet
$1.5 \leq I_i \leq 1.99$	Very wet

$1.0 \leq I_i \leq 1.49$	Moderately wet
$-0.99 \leq I_i \leq 0.99$	Close to normal
$-1.0 \leq I_i \leq -1.49$	Moderately dry
$-1.5 \leq I_i \leq -1.99$	Very dry
$I_i \leq -2$	Extremely dry

I_i is the standardized flow index, Q_i is the annual flow of a particular year, and $\overline{Q}$ is the annual flow average over the period; σ is the standard deviation during the time.

The indices outlined in Table 3 offer critical insights into the extreme rainfall and river flow patterns. These indicators are vital for evaluating flood risks, improving water resource management, and understanding the hydrological dynamics of rivers in the region. Such analyses are indispensable for detecting trends in extreme events, supporting effective decision-making and planning to address climate variability and its effects on water resources.

Table 3: Indices for extreme rainfall and river flow.

Index	Name of the Index	Index Description	Units
SDII	Simple daily rainfall index	Ratio of yearly total rainfall to the number of rainy days	mm/day
RX1DAY	Max 1-day rainfall	Highest single-day rainfall total recorded within a year	mm
RX5DAY	Max 5-day rainfall	Highest total rainfall over any 5 consecutive days in a year	mm
R95P	Very wet days	Total yearly rainfall exceeding the 95th percentile (1991–2020)	mm
R99P	Extremely wet day	Total yearly rainfall exceeding the 99th percentile (1991–2020)	mm
Qmax	Peak discharge	Highest yearly river discharge (1991–2020)	m^3/s
Q95P	High-flow days	Total yearly streamflow from days exceeding the 95th percentile (1991–2020)	m^3/s
Q99P	Very-high-flow days	Total yearly streamflow from days exceeding the 99th percentile (1991–2020)	m^3/s

- Trend and change-point detection
  - Tests for Trend Analysis

This study utilized the modified Mann–Kendall (MMK) test to examine the spatial and interannual trends in extreme rainfall and river flow across the BRB and its upper catchment

during the period 1991–2020. The MMK test, a widely recognized nonparametric approach, was chosen for its suitability in analyzing the 30-year data series. It is particularly effective for hydrological and climatic datasets due to its minimal reliance on assumptions about the data distribution (Hamed and Ramachandra Rao, 1998; Ndiaye et al., 2023; Yue and Wang, 2002). The use of MMK instead of the traditional Mann–Kendall (MK) test is justified by the presence of autocorrelation in the time series data, which violates the independence assumption of the MK test. As highlighted in (Ehsanzadeh and Adamowski, 2010), autocorrelation can significantly impact trend detection, with positive autocorrelation artificially inflating the significance of trends and negative autocorrelation masking actual trends.

An enhancement of the traditional Mann–Kendall (MK) test, the MMK method addresses the issue of autocorrelation in time series data, which can distort trend detection. Positive autocorrelation can exaggerate trend estimates, while negative autocorrelation can mask actual trends (Hamed and Ramachandra Rao, 1998). By adjusting the variance of the MK statistic to account for autocorrelation, the MMK test minimizes the likelihood of false trend detection (Kundzewicz and Robson, 2004). The preprocessing steps involved in applying the MMK test, such as removing seasonal patterns, were necessary to isolate long-term trends from periodic fluctuations. However, these adjustments can affect the statistical significance of indices sensitive to short-term variability, such as the R95P (Hamitouche et al., 2024).

The MMK test functions similarly to the classical MK test but incorporates adjustments to the variance of the test statistic (S) to account for autocorrelation effects. This makes it particularly effective for datasets with complexities such as tied values, seasonality, or missing data. Although robust, the MMK test requires additional preprocessing steps, including addressing tied values and seasonal patterns, which add complexity to its implementation (Burn and Hag Elnur, 2002; Yue and Wang, 2002). It has been demonstrated that methods accounting for persistence in time series, such as MMK, are crucial for ensuring reliable trend detection in data prone to autocorrelation (Serinaldi and Kilsby, 2016).

$$\mathrm{Var}\,(s) = \frac{1}{18}\big(n(n-1)(2n+5)\big)\frac{n}{ns^*} \tag{6}$$

The parameter ns* is utilized to correct the effective number of observations, considering the autocorrelation in the data.

$$\frac{n}{ns^*} = 1 + \frac{2}{n(n-1)(n-2)} \sum_{s=1}^{m} (n-s) \quad (7)$$

$\frac{n}{ns*}$ is a correction factor due to the autocorrelation present in the data, m représente le nombre de blocs ou de groupes dans lesquels la série de données est divisée.

The trend analysis was conducted using the "modifiedmk" package within the R programming environment, which is specifically designed to apply the modified Mann–Kendall (MMK) test while accounting for autocorrelation in time series data. This test evaluates the null hypothesis (H0) of "no trend" against the alternative hypothesis (H1) indicating the presence of a trend. A 5% significance level was employed to determine the statistical significance of the identified trends, aligning with common practices in hydrological and climatic research (Sang et al., 2014).

- Change-point detection tests

Pettitt's test, a nonparametric method introduced in (Pettitt, 1979), is widely used to detect abrupt changes in time series data, such as shifts in climate or hydrological records. This test identifies statistically significant changes in the mean of a dataset by evaluating the null hypothesis (H0) of uniformity against the alternative hypothesis (H1) that a change has occurred. A change point is deemed statistically significant if the *p*-value is ≤0.05. Pettitt's test has been extensively applied in studies to analyze shifts in climatic and hydrological conditions (Khaliq et al., 2009; Pettitt, 1979).

The standard normal homogeneity test (SNHT) is a robust statistical tool for detecting discontinuities in time series, particularly at the beginning or end of a dataset. It is effective even with missing values, making it reliable for climatic and hydrological studies. The test's strength lies in its ability to identify structural breaks in the data through its sensitivity and straightforward implementation. The application of the SNHT relies on the utilization of the following equation:

$$Q_i = Y_i - \frac{\sum_{j=1}^{k} \rho_j{}^2 x_{ij} \frac{\bar{y}}{\bar{x}}}{\sum_{j=1}^{k} \rho_j{}^2} \quad (8)$$

$$1 \leq i \leq n \text{ and } 1 \leq j \leq k$$

The base series comprises values denoted as Y_i for each year i, while the reference series, labeled j, contains observations represented as X_{ij} for each year i. The value ρ_j de-notes the correlation coefficient between the base series and the reference series j. $\overline{X}$ et $\overline{Y}$ represents the mean (average) of the values in the dataset

These tests, the MMK, Pettitt's test, and the SNHT, as shown in Figure 6, are valuable tools for trend detection and the identification of structural changes. However, their effectiveness and relevance depend on the specific characteristics of the analyzed data

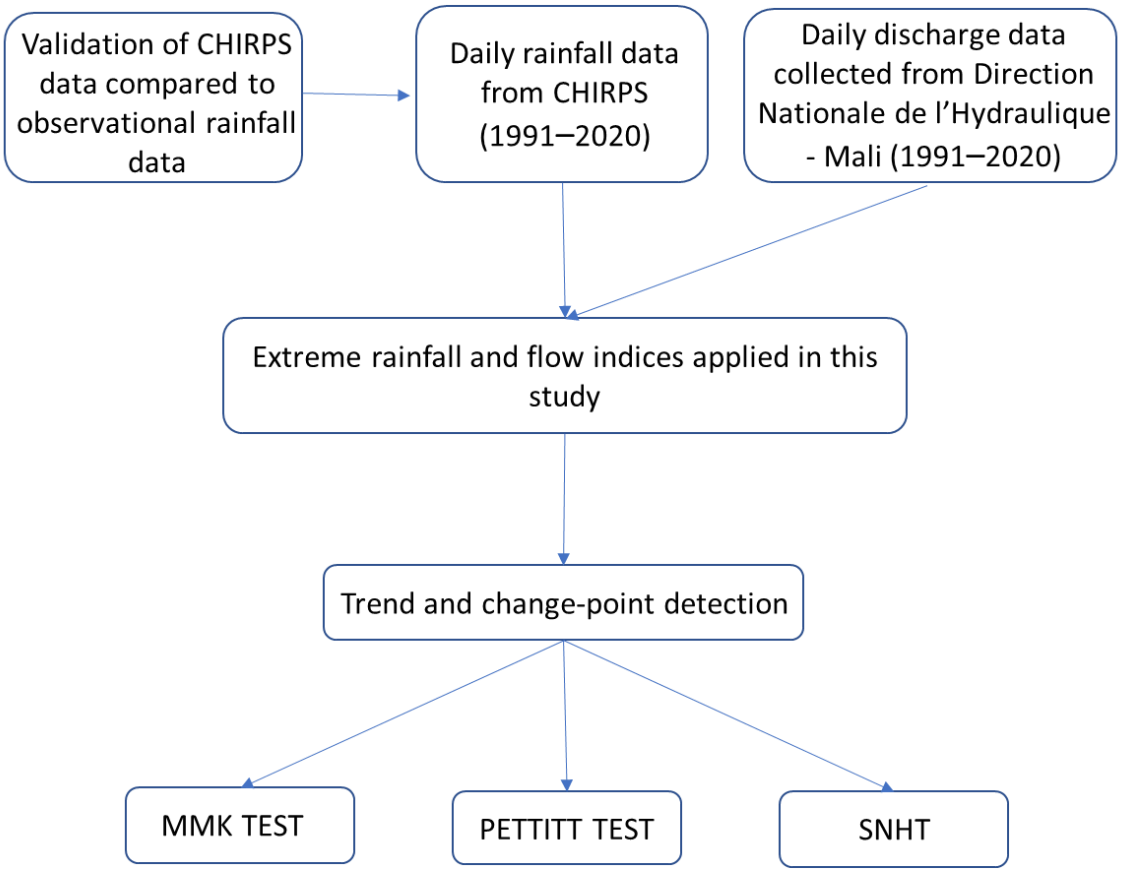

Figure 6: Flowchart illustrating the methodology applied.

- Extreme indices calculation

Extreme rainfall indices such as SDII, RX1DAY, RX5DAY, R95P, and R99P were calculated to capture the intensity and frequency of extreme precipitation events. For river discharge, indices like Qmax, Q95P, Q99P, and the Standardized Flow Index (SFI) were computed. These indices provide critical insights into hydrological extremes, such as floods and droughts, and their potential impacts on water resource management. The Standardized Flow Index (SFI) was calculated using the following formula:

$$SFI = \frac{X - \mu}{\sigma} \tag{9}$$

Here, X represents the observed flow, μ is the mean flow, and σ is the standard deviation of the flow. The calculated indices are crucial for understanding trends and variability in extreme hydrological events, providing valuable information for decision-making and planning.

3.3. Machine learning models for water fluctuation prediction for the third specific objective

In our research, we focused on a comprehensive dataset incorporating various environmental parameters such as total evaporation, water levels for all the available gauge stations (e.g., Pankourou, Sofara, Bougouni, Douna, and Mopti), and precipitation levels. The dataset spans from 1983 to 2020, enabling to analyze trends and changes in hydrological patterns over time. The input features included 'Tot_evap', 'WL_Sof', 'WL_Bou', 'WL_Pank', 'P', 'WL_Dou', and 'WTD', while 'WL_Mop' served as the target variable. We maintained the chronological integrity of the data by splitting it sequentially. In the first cross-validation fold (Figure 7a), we allocated the first 60% of the data for training, the next 20% for testing, and the final 20% for validation. In the second fold (Figure 7b), the first 20% was used for testing, the middle 60% for training, and the last 20% for validation. This input and output data configuration process allowed us to train and evaluate our models on different temporal segments, enhancing their robustness and predictive reliability in modeling complex hydrological patterns.

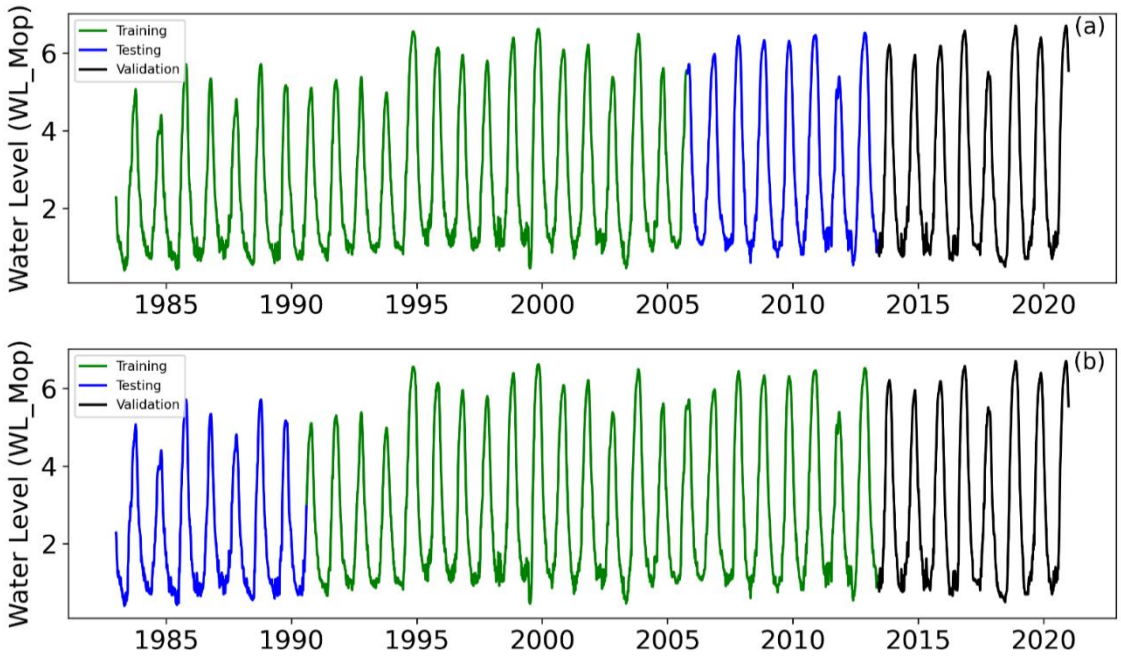

Figure 7: Water level time series for Fold 1 (a), and Fold 2 (b).

- Model selection and algorithms

To predict water level fluctuations, four machine learning algorithms were employed: XGB, SVM, RF and GB. These algorithms were chosen for their ability to handle non-linear relationships and interactions between multiple variables. XGB, for instance, is highly efficient in handling large datasets with complex structures, while SVM is well-suited for managing non-linear data. RF and GB, both ensemble methods, are effective in improving prediction accuracy by combining multiple decision trees. This diverse set of algorithms ensured a robust analysis of water fluctuation patterns.

- RF is an ensemble learning method primarily used for classification and regression. It builds multiple decision trees and merges them together to get a more accurate and stable prediction. RF adds additional randomness to the model, while growing the trees. Instead of searching for the most important feature while splitting a node, it searches for the best feature among a random subset of features. This results in a wide diversity that generally results in a better model.

$$Y = \frac{1}{n}\sum_{i=1}^{n} T_i(X) \tag{10}$$

Where T_i represents each decision tree in the forest, X is the input vector, and n is the number of trees.

- GB constructs a predictive model in the form of an ensemble of weak prediction models, typically decision trees. It builds the model in a stage-wise fashion and generalizes them by allowing optimization of an arbitrary differentiable loss function.

$$F_{t+1}(x) = F_t(x) + \gamma_t h_t(x) \tag{11}$$

Where F_t is the model at iteration t, $h_t(x)$ is the weak learner added at the t-th stage, and γ_t is the learning rate at stage t.

- XGB is an implementation of gradient boosting machines that is designed to be highly efficient, flexible, and portable. XGB improves on the base GBM framework through systems optimization and algorithmic enhancements.

$$\theta_{t+1} = \theta_t - \eta \sum_{i=1}^{n} g_i h(x_i, \theta_t) \tag{12}$$

Where θ represents the model parameters, η is the learning rate, g_i are the gradients, and h is the Hessian matrix of the loss function.

- SVM is a powerful and versatile classification technique that works both on linear and non-linear data. It creates a hyperplane or set of hyperplanes in a high or infinite-dimensional space, which can be used for classification, regression, or other tasks.

$$\min\left(\frac{1}{2}\omega^{\top}\omega + C\sum_{i=1}^{n}\zeta_i\right) \tag{13}$$

Subject to the constraints:

$$y_i(\omega^{\top}\phi(x_i) + b) \geq 1 - \zeta_i, \quad \text{for each i}$$

Here, ω is the normal vector to the hyperplane, b is the bias term, C is the regularization parameter, ζ_i are the slack variables, ϕ maps input data to a higher-dimensional space, and y_i are the class labels.

- Data partitioning and validation

The dataset was divided into training, testing, and validation subsets using a chronological split to preserve temporal dependencies. In the first fold, 60% of the data was used for training, 20% for testing, and 20% for validation. In the second fold, 20% of the data was used for testing, 60% for training, and the remaining 20% for validation. This approach ensured that the models were tested on different temporal segments, enhancing their generalizability. Hyperparameter optimization was conducted during training to improve model performance. The validation process helped in identifying the best-performing model for predicting water level fluctuations.

- Model evaluation metrics

The performance of the machine learning models was evaluated using three key metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2). RMSE measures the average magnitude of prediction errors, providing insights into the accuracy of the models. MAE quantifies the average absolute difference between observed and predicted values, while R^2 assesses how well the model explains the variability in the observed data. Together, these metrics provided a comprehensive evaluation of the models' predictive capabilities.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|O_i - P_i| \tag{14}$$

MAE measures the average magnitude of the errors between predicted values (P_i) and observed values (O_i), without considering the direction of the errors (i.e., whether they are over or underestimations).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \tag{15}$$

R^2 quantifies how well the variance in the observed data can be explained by the model. It is a statistical measure that shows the proportion of the variance for a dependent variable that's predicted from the independent variables.

3.4. Future hydroclimatic projections for the fourth specific objective

The study utilized extensive hydrological and meteorological data collected since 1983. This included daily river level at Mopti and Digital Elevation Model (DEM) data were obtained from the Copernicus website (Table 4). The limited number of ground-based precipitation stations in the BRB highlights the value of the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) dataset (Funk et al., 2015). CHIRPS, widely used for modeling, has proven reliable in various studies (Basse et al., 2021; Du et al., 2024; Fofana et al., 2022; Kouyaté et al., 2025; Oyerinde et al., 2017; Rameshwaran et al., 2021).

Table 4: Data used in this study including observed river water levels.

Data	Source	Time Series
CHIRPS	IRI/LDEO Climate Data Library	1983 – 2020
DEM	USGS/NASA	-
River Water Level	National Hydraulic Directorate of Mali	1983 – 2020

Hence, applying both a mid-range mitigation scenario (SSP2-4.5) and a higher-emission pathway (SSP5-8.5) captures the broad spectrum of potential greenhouse gas trajectories in West Africa (Almazroui et al., 2020; Limantol et al., 2023; Sawadogo et al., 2024), aligning with established climate modeling frameworks and reflecting the region's pronounced vulnerability to extreme hydroclimatic events. In addition, Table 5 presents data from 10 CMIP6 outputs for daily precipitation and temperature spanning from 1983 to 2050, providing critical inputs for future projections.

Table 5: List of CMIP6 models used in the study with their descriptions and spatial resolutions.

Model	Description	Spatial resolution (lon × lat)
ACCESS-CM2	Commonwealth Scientific and Industrial Research Organisation (Australia)	192 × 144
ACCESS-ESM1-5	Commonwealth Scientific and Industrial Research Organisation (Australia)	192 × 145
BCC-CSM2-MR	Beijing Climate Center, Beijing (China)	320 × 160
CanESM5	Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada (Canada)	128 × 64
GISS-E2-1-G	Goddard Institute for Space Studies, National Aeronautics and Space Administration (USA)	144 × 90
MIROC6	Japan Agency for Marine-Earth Science and Technology (Japan)	256 × 128
MPI-ESM1-2-HR	Max Planck Institute for Meteorology (Germany)	384 × 192
MPI-ESM1-2-LR	Max Planck Institute for Meteorology (Germany)	192 × 96
NESM3	Nanjing University of Information Science and Technology (China)	192 × 96
TaiESM1	Seoul National University (Republic of Korea)	288 × 192

- Methodology

The methodology begins with data collection, where historical and future climate data are gathered and preprocessed to ensure consistency and compatibility. Historical rainfall data from the CHIRPS dataset (1983–2020) serves as the reference for calibrating and validating the climate models. CHIRPS data is complemented by DEM data for Topographic Wetness Index (TWI) calculation, a critical input for understanding soil moisture and water flow dynamics. For future projections, daily rainfall and temperature data from 10 CMIP6 models (Table 5) were extracted under SSP2-4.5 and SSP5-8.5 scenarios, capturing both moderate and extreme climate pathways.

The second component involves model selection and training. An ensemble mean of the CMIP6 model outputs is computed for rainfall and temperature, followed by an evaluation using a Taylor diagram to compare the performance of individual models against reference. The best-performing model is then selected for use in future simulations. ML techniques, specifically RF and GB, are employed to predict future water levels based on rainfall, temperature, and TWI inputs. The dataset is split into training and testing subsets using a two-fold cross-validation approach. Before cross-validation, the last 20% of the dataset is removed as an unseen

validation set. The remaining 80% is divided into two folds: in the first fold, 20% is used for testing and the remaining 60% for training, while in the second fold, the first 60% is used for training and the remaining 20% for testing. Hyperparameter optimization is conducted during this process, and the training hyperparameter models are retrained on the full 80% dataset and validated using the unseen 20% dataset to ensure unbiased error assessment.

The final component focuses on characterizing hydroclimatic events. Rainfall and water level projections for 2021–2050 are analyzed to compute key hydroclimatic indices (Table 6).

Table 6: Description of hydroclimatic indices used in the study.

Index	Name of the Index	Index Description	Units
CDD	Consecutive dry days	Maximum number of consecutive days with daily rainfall < 1 mm	days
CWD	Consecutive wet days	Maximum number of consecutive days with daily rainfall ≥ 1 mm	days
RX1DAY	Max 1-day rainfall	Highest single-day rainfall total recorded within a year	mm
WLmax	Peak water level	Highest yearly river water level (2021–2050)	mm

These indices are important for understanding drought risks, prolonged wet periods, and flood potential. Statistical tests are applied to these indices to detect trends and changes. The Pettitt test is used to identify abrupt changes or breakpoints in the time series, while the Modified Mann-Kendall (MMK) test is employed to detect long-term trends, accounting for serial correlation in the data. This integrated approach summarize in the Figure 8 ensures a comprehensive understanding of future hydroclimatic conditions in the BRB. By combining data collection methods, climate scenario modeling, and ML, the methodology provides robust predictions and insights into hydroclimatic trends.

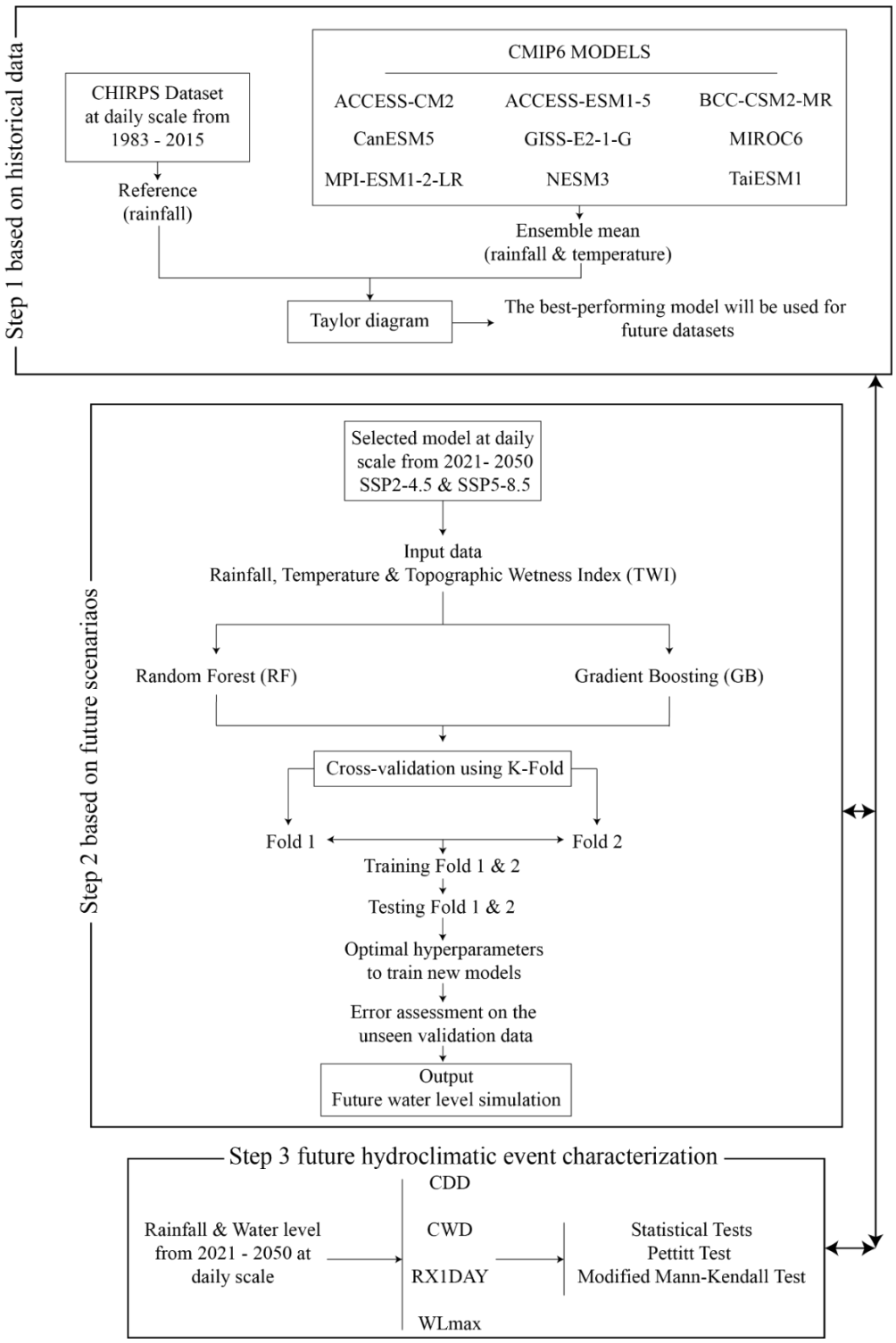

Figure 8: Methodology summarized in the flowchart.

- Model selection and training

The ensemble mean of CMIP6 outputs for rainfall and temperature was computed to improve the reliability of future projections. Model performance was evaluated using a Taylor diagram,

which compares the correlation, standard deviation, and centered RMSE of each model against observations. The best-performing model was selected for simulating future hydroclimatic conditions. Machine learning techniques, specifically RF and GB, were used to predict future water levels based on rainfall, temperature, and TWI inputs. A two-fold cross-validation framework was implemented to enhance model robustness and prevent overfitting.

- Hydroclimatic indices

Key hydroclimatic indices, including CDD (Consecutive Dry Days), CWD (Consecutive Wet Days), RX1DAY (Maximum 1-Day Rainfall), and WLmax (Peak Water Level), were analyzed for the period 2021–2050. These indices provide valuable insights into drought risks, prolonged wet periods, and potential flood events. Statistical tests were applied to these indices to detect trends and changes, offering a detailed understanding of future hydroclimatic patterns in the BRB.

$$\mathrm{CDD} = \max(n: P_i < 1\ \mathrm{mm\ for\ i} = 1,2,\dots,n) \quad (16)$$

$$\mathrm{CWD} = \max(n: P_i \geq 1\ \mathrm{mm\ for\ i} = 1,2,\dots,n) \quad (17)$$

$$\mathrm{RX1DAY} = \max(P_i: P_i\ \mathrm{for\ each\ day\ i}) \quad (18)$$

$$\mathrm{WLmax} = \max(L_i: L_i\ \mathrm{for\ each\ observation\ period\ i}) \quad (19)$$

n represents the number of consecutive days in the sequence being considered. For CDD and CWD, it signifies the total number of days in the longest run of days meeting certain precipitation criteria. Pi stands for the precipitation on day i. It is used in all indices to represent daily precipitation amounts. It is compared against a threshold (1 mm) to determine whether a day is dry or wet. In RX1DAY, it's used to find the maximum precipitation recorded in any single day. Li represents the water level on observation period i. Li is used to find the highest water level recorded during the observation periods.

- Statistical analysis

To detect trends and changes in hydroclimatic indices, Pettitt's Test and the Modified Mann–Kendall (MMK) test were applied. Pettitt's Test identified abrupt changes or breakpoints in the time series, while the MMK test accounted for serial correlation to detect long-term trends. These analyses provided a comprehensive assessment of potential changes in hydroclimatic conditions under different climate scenarios.

Figure 9 provides a visual summary of the methodology adopted in this thesis, outlining the data collection and analysis processes across the four primary objectives.

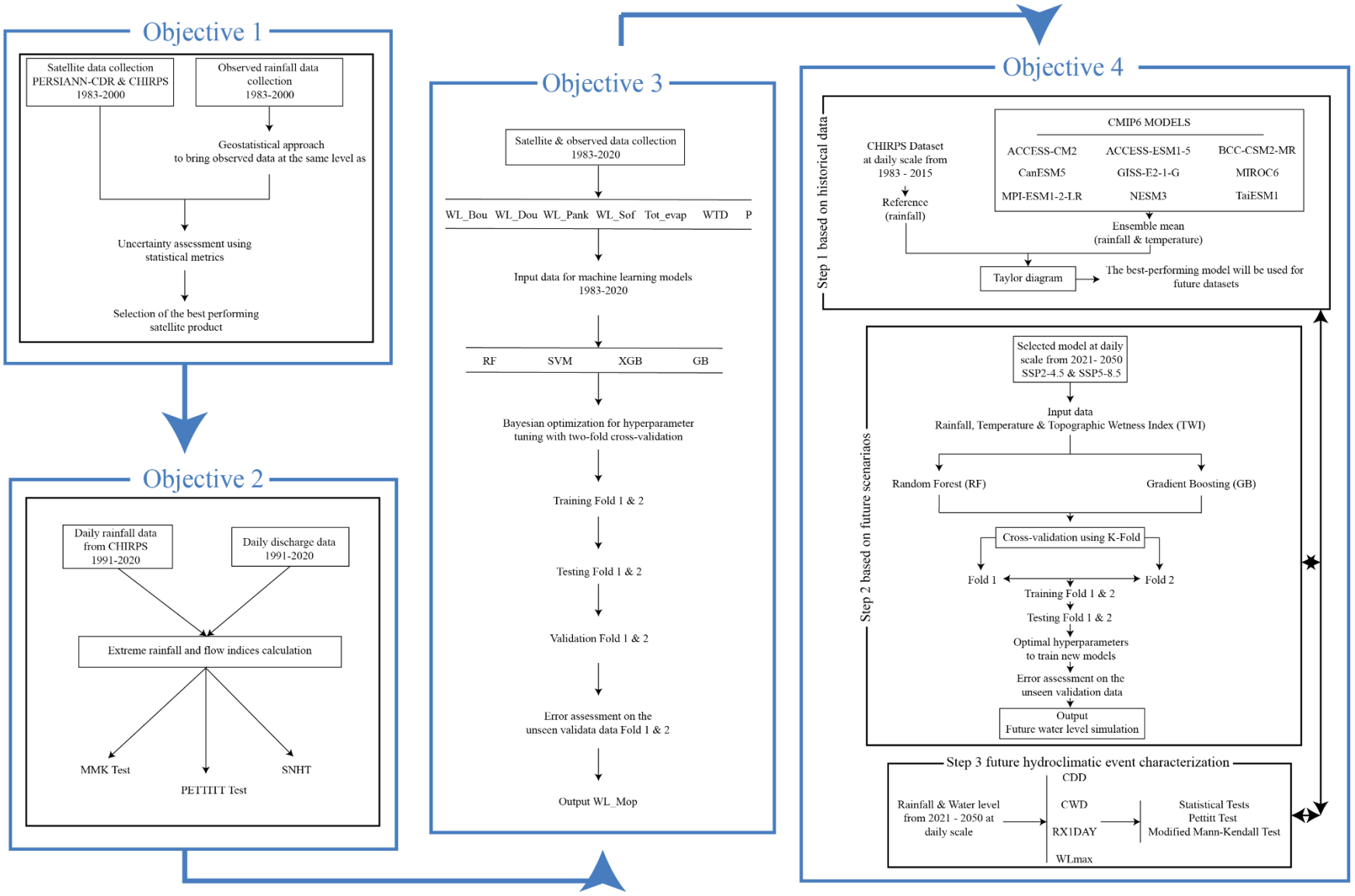

Figure 9: Thesis methodology overview.

3.5. Partial conclusion

Chapter 3 has detailed a multifaceted research methodology designed to address the study's four key objectives. The chapter began by constructing a robust Reference Rainfall Product (RRP) through geostatistical methods and evaluating satellite-derived precipitation datasets (CHIRPS and PERSIANN-CDR). These preliminary steps established a reliable foundation for subsequent hydrological analyses. The chapter then underscored the importance of identifying and characterizing extreme hydroclimatic conditions both in rainfall and river discharge by employing standardized indices. This facilitated a systematic approach to understanding flood- and drought-related risks in the BRB.

A critical portion of this chapter involved implementing and assessing four machine learning algorithms (XGB, SVM, RF, and GB) to predict river water-level fluctuations, highlighting the role of data-driven modeling in capturing complex, non-linear hydrological behaviors. Finally, the future hydroclimatic projections were generated under two CMIP6 scenarios (SSP2-4.5 and SSP5-8.5), showcasing how historical data and modern modeling frameworks can be combined to anticipate climate impacts on river basins. In essence, this chapter provides the necessary methodological rigor ranging from geostatistical validation to advanced computational modeling to ensure that the subsequent results and discussions are both reliable and contextually grounded in the evolving hydroclimatic realities of the region.

Chapter 4: Satellite products assessment and Hydroclimatic information for the BRB

West Africa faces growing challenges related to climate variability and human-induced changes, which significantly impact its river systems. Among these challenges are alterations in rainfall patterns and river flows, driven by factors such as dam construction, land-use changes, and global climate change (Oyerinde et al., 2017; Sylla et al., 2018). While research on West Africa's river systems has primarily focused on larger and better-documented basins such as the Senegal and Niger, the BRB has received comparatively less attention, despite its critical role as a tributary of the Niger River and its unique hydrological vulnerabilities (Dos Santos, 2023). Understanding and quantifying these trends is crucial for ensuring sustainable water resource management and mitigating the effects of hydrological extremes (Descroix et al., 2012). The hydroclimatic variability in West Africa has been extensively documented, with research highlighting significant temporal and spatial changes in rainfall and river flow across the region. Studies have reported an increase in extreme rainfall events and river discharge in the Senegal River Basin, with a notable shift to wetter conditions (Ndiaye et al., 2023). Significant transitions in annual rainfall patterns have also been identified, with periods of drought followed by partial recovery after 1994 (Bodian et al., 2020). However, the BRB faces specific challenges due to its semi-arid environment, reliance on rain-fed agriculture, and significant human pressures from land-use changes and water infrastructure development. Hydrological models have been used to assess the basin's response to rainfall variability, identifying substantial impacts on river flow and water resource availability (Louvet et al., 2016). The intricate interactions between surface water and groundwater in the basin have provided critical insights into its hydrological processes (Mahe, 2009). This study also contributes to addressing these knowledge gaps by focusing specifically on the rainfall and river flow trends in the BRB, leveraging statistical methods that provide a detailed temporal analysis.

Despite these advances, the BRB remains less studied than major basins like the Senegal and Niger. This study aims to fill this gap by conducting a comprehensive assessment of the trends and significant changes in rainfall and river flow within the BRB. By integrating historical

observations and statistical techniques, this research seeks to provide a deeper understanding of the basin's hydroclimatic dynamics.

A significant portion of this chapter has been published in a peer-reviewed journal: Kouyaté, F., Guédjé, F.K., Ndiaye, A., Ganni Mampo, O.M., 2025. Spatial and Temporal Variability of Extreme Hydroclimatic Events in the BRB. Hydrology 12, 5. https://doi.org/10.3390/hydrology12010005.

4.1. Structural analysis

The variographic analysis involved an exponential type variogram built from observed rainfall data. This variogram is oriented in an east-west direction, indicating a defined structure. Along this direction, the experimental variogram can be fitted to an exponential model. The range is estimated at 1.8° (equivalent to 199.8 km) in the east-west direction, with a sill reaching 78 km. The plateau is reached beyond our estimated range of 1.8°, meaning that there is no longer a dependency between the data beyond this distance. This could likely be due to sampling errors or the intrinsic quality of the data (Bostan et al., 2012; Katipoğlu, 2022). Subsequently, the reference rainfall product was calculated using ordinary kriging at the same spatial resolution (0.25° by 0.25°) and at the same grid points as the satellite products. The parameters from the variographic analysis were employed to construct the rainfall fields for the BRB.

4.2. Statistical analysis of precipitation data

A comparative analysis of the CHIRPS (Table 7) and PERSIANN-CDR (Table 8) products based on statistical metrics such as Mean Absolute Error (MAE), Normalized Root Mean Square Error (NRMSE), and Percent Bias (PBIAS) reveals distinct performance characteristics for each dataset over the years 1983 to 2000. CHIRPS consistently shows lower MAE and NRMSE values across the years compared to PERSIANN-CDR, indicating that CHIRPS generally has better accuracy in estimating precipitation. For example, in 1983, CHIRPS recorded an MAE of 2.19 mm and an NRMSE of 94.1, significantly lower than PERSIANN-CDR's 3.26 mm and 120.6, respectively. This pattern is consistent through the years, with CHIRPS maintaining lower error rates. PBIAS, which measures the tendency of the models to

overestimate or underestimate precipitation, also shows significant differences. CHIRPS tends to have a much lower and occasionally negative PBIAS, suggesting a more balanced estimation, whereas PERSIANN-CDR exhibits a higher PBIAS, particularly evident in years like 1990 and 1991, where it reaches as high as 67% and 73.3%, respectively. Overall, CHIRPS demonstrate more reliable and accurate precipitation estimates than PERSIANN-CDR, with lower error metrics and a more neutral bias, making it potentially more suitable for applications requiring high precision in rainfall data. However, the higher PBIAS in PERSIANN-CDR might make it useful in specific scenarios where consistent overestimation is predictable and can be calibrated accordingly.

Table 7: Statistical assessment parameters of the CHIRPS product.

Year	MAE (mm)	NRMSE	PBIAS (%)
1983	2.19	94.1	-2.6
1984	2.2	100.1	4.4
1985	2.38	97.0	-5.5
1986	2.46	90.4	-9.8
1987	2.16	88.1	-2.6
1988	2.68	84.3	-14.1
1989	3.23	94.1	-30.7
1990	2.62	93.1	4.9
1991	2.74	96.2	-2.8
1992	3.04	105.8	-5.3
1993	2.16	81.8	1.1
1994	3.12	81.8	-15
1995	2.54	88.3	-4.1
1996	2.22	90.1	2.3
1997	2.92	94.1	-12.9

1998	2.95	112.0	13.4
1999	2.68	85.5	-3.8
2000	2.6	110.0	16.0

Table 8: Statistical assessment parameters of the PERSIANN-CDR product.

Year	MAE (mm)	NRMSE	PBIAS (%)
1983	3.26	120.6	47.8
1984	3.16	125.6	9.6
1985	3.46	112.7	52.0
1986	3.75	113.3	28.2
1987	3.3	108.0	52.4
1988	4.39	109.3	23.6
1989	4.01	100.6	8.6
1990	3.45	96.1	67.0
1991	4.54	121.3	73.3
1992	3.72	108.5	42.1
1993	3.52	122.9	51.9
1994	4.64	108.7	13.0
1995	3.86	112.1	57.0
1996	4.08	129.1	75.4
1997	4.29	109.4	37.9
1998	3.57	114.3	46.8
1999	3.75	107.6	35.4
2000	3.32	119.8	55.4

4.3. Assessment of rainfall satellite products

We compared CHIRPS and PERSIANN-CDR with ground-based reference rainfall product (RRP) data over the BRB from 1983 to 2000 to determine which rainfall estimate provides more reliable information over BRB. The RRP has been created through rainfall ground-based measurements and geostatistical interpolation method. The estimates of average daily precipitation in Figure 10 highlight significant divergences between CHIRPS and PERSIANN-CDR. As seen, CHIRPS mimics the RRP data more consistently, except in the southeast of the basin (Figure 11), towards Côte d'Ivoire and Burkina Faso, where an underestimation is visible. PERSIANN-CDR, on the other hand, displays a generalized overestimation except in a central area of the basin. The quantitative results, detailed in Figure 10a-c, support these findings.

Furthermore, the Quantile Mapping bias-correction analysis (Figure 10d–e) demonstrates a clear improvement in the agreement between satellite-derived rainfall estimates and the RRP, particularly for CHIRPS, which shows higher post-correction consistency and reduced bias.

We further calculated error evaluation statistics (Figure 12a-b) such as MAE, NRMSE, and PBIAS. CHIRPS generally presented the lowest estimation errors, while PERSIANN-CDR recorded the highest. However, it should be noted that for some years, such as 1994, CHIRPS had a higher NRMSE than PERSIANN-CDR, indicating that the accuracy of these products can vary from year to year. CHIRPS tends to slightly underestimate except during four specific years where overestimations are observed. In contrast, PERSIANN-CDR showed a consistent trend of overestimation throughout the study period. This characteristic of overestimation is crucial to consider for water resource management and flood mitigation strategies, as it could lead to incorrect forecasts of hydrological events.

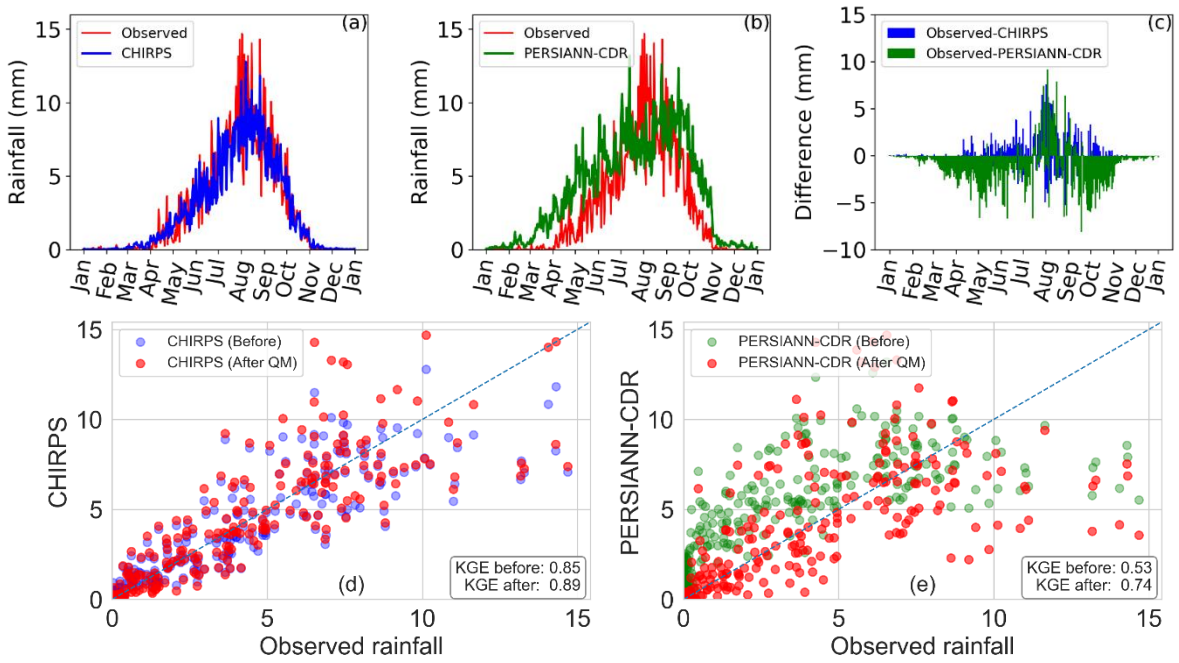

Figure 10: Daily interannual seasonal cycles of CHIRPS (a) and PERSIANN-CDR (b) relative to the RRP (c) from 1983 to 2000, and corresponding Quantile Mapping (QM) bias-correction results for CHIRPS (d) and PERSIANN-CDR (e) compared to observed rainfall. The scatter plots illustrate improvements in distribution alignment and Kling-Gupta Efficiency (KGE) after bias correction.

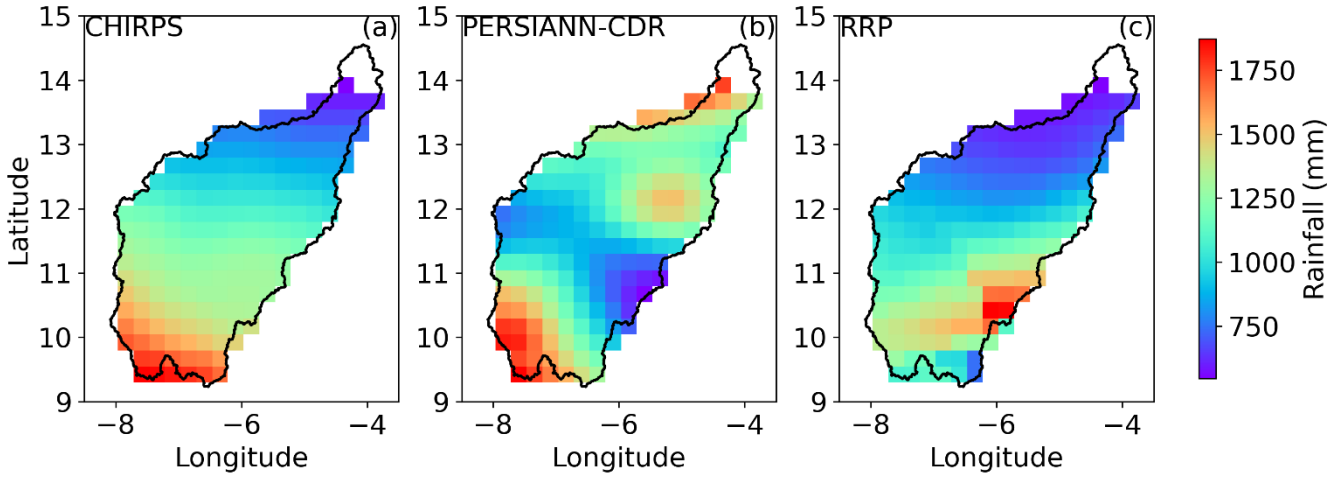

Figure 11: Spatial variation of annual rainfall in the Bani River basin estimated by CHIRPS (a) and PERSIANN-CDR (b) from 1983 to 2000. Annual reference rainfall product (RRP) from 1983 to 2000 (c).

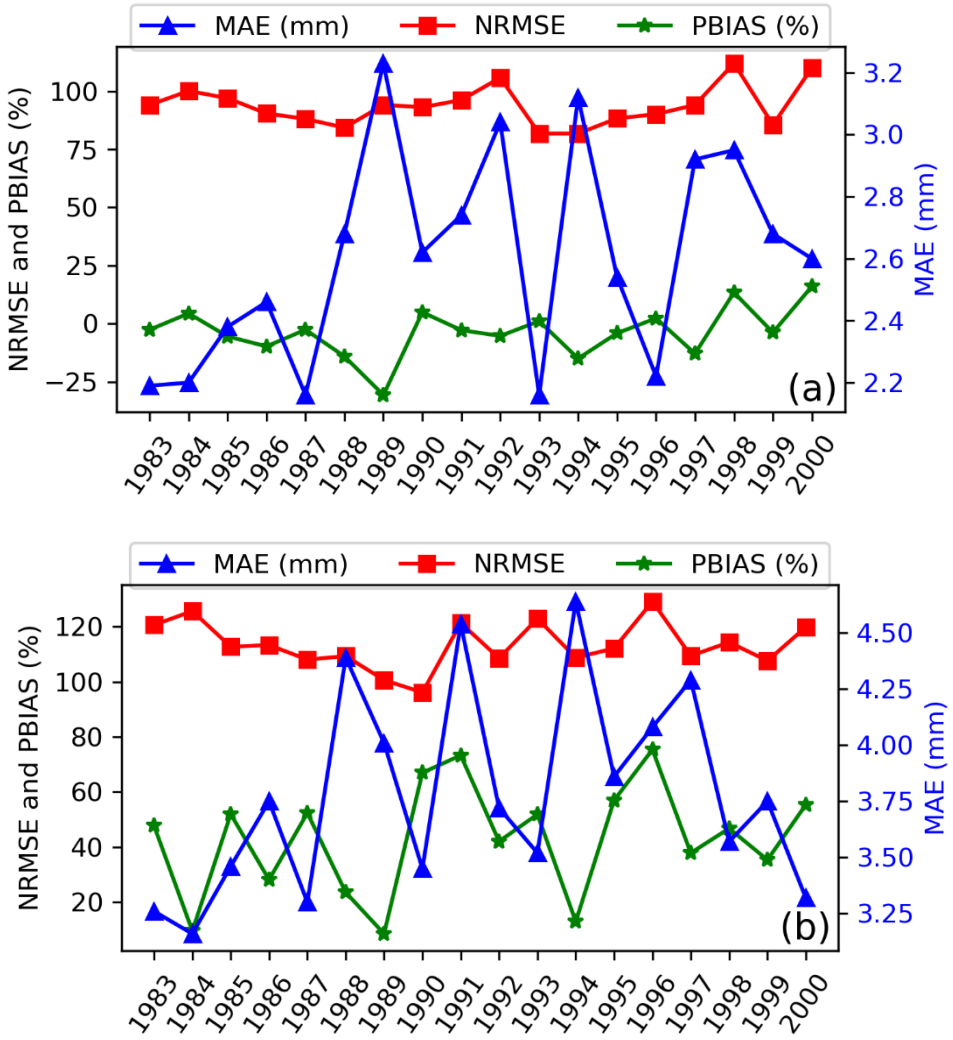

Figure 12: Statistical metrics for the comparison of CHIRPS (a) and PERSIANN-CDR (b) with RRP.

The results of a comparative analysis of satellite rainfall estimation products suggested that CHIRPS provide reliable precipitation estimates in the basin. This is attributed to the integration of ground observations with satellite data in CHIRPS dataset, which enhances its spatial accuracy and temporal resolution as the critical features for managing the hydrological dynamics of a region with significant climatic variability. The result is supported by recent publications such as those by Houngnibo et al. (2023), which validate the effectiveness of CHIRPS in similar geographic settings (Basse et al., 2021; Du et al., 2024; Fofana et al., 2022; Kouyaté et al., 2025; Oyerinde et al., 2017; Rameshwaran et al., 2021). This helps with filling the gaps caused by the observational rainfall data scarcity and the non-availability of

conventional data, ensuring that the hydrological authority in the region can maintain water management operations despite the challenging conditions.

4.4. Assessment of CHIRPS station data quality

The NSE values presented in this study (Figure 13), ranging from 0.95 to 0.98, underscore the strong agreement between the CHIRPS satellite-based rainfall estimates and the ground-based station observations across the four locations. These high NSE values validate CHIRPS as a reliable tool for accurately estimating rainfall in the BRB region.

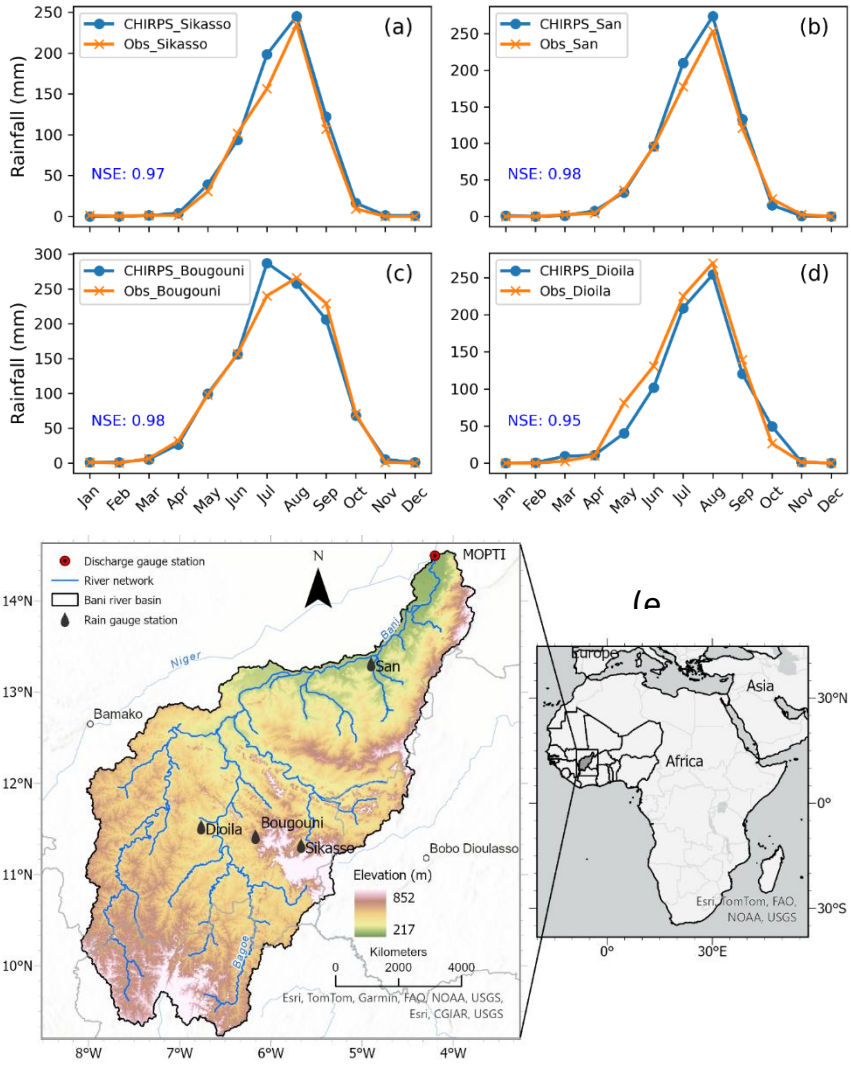

Figure 13: Monthly rainfall comparison between CHIRPS data and observed station measurements in the BRB (2011–2019): (a) Segou, (b) San, (c) Bougouni and (d) Dioila; location of rain gauge station (e).

The observed trends in Figure 13a–d demonstrate that CHIRPS effectively captures the temporal and seasonal rainfall dynamics, particularly during the rainy season (June to September). We were able to collect the recent observation data from the complete series of 4 stations distributed throughout the basin, as shown in figure 8e. Minor discrepancies in some months, such as the peak rainfall periods in Figure 7d, are minimal and fall within the acceptable limits for satellite-based estimations. These findings align with prior research. For instance, ref. (Funk et al., 2015) highlighted that CHIRPS is a dependable dataset for monitoring rainfall, particularly in regions with sparse observational networks. Similarly, ref. (Dembélé and Zwart, 2016) validated CHIRPS in West Africa, demonstrating strong correlations with station data, making it highly suitable for agricultural planning.

4.5. Spatial variation in extreme rainfall indices over the BRB

Figure 14 illustrates the spatial variation of the extreme rainfall indices, the RX1DAY, RX5DAY, R95P, R99P and SDII, for the period 1991–2020 during the rainy season (June to September) across the study area. Higher values of these indices are concentrated in the southern regions, with a maxima of approximately 10 mm/day for the SDII, 54 mm for the RX1DAY, 96 mm for the RX5DAY, 20 mm for the R95P, and 36 mm for the R99P. These patterns highlight a higher intensity and frequency of extreme rainfall in the southern part of the basin, while the lower values in the northern regions reflect reduced exposure to extreme rainfall events upstream. The coordinate system used for this spatial analysis is the WGS84 geographic coordinate system, ensuring accurate and reproducible spatial representation.

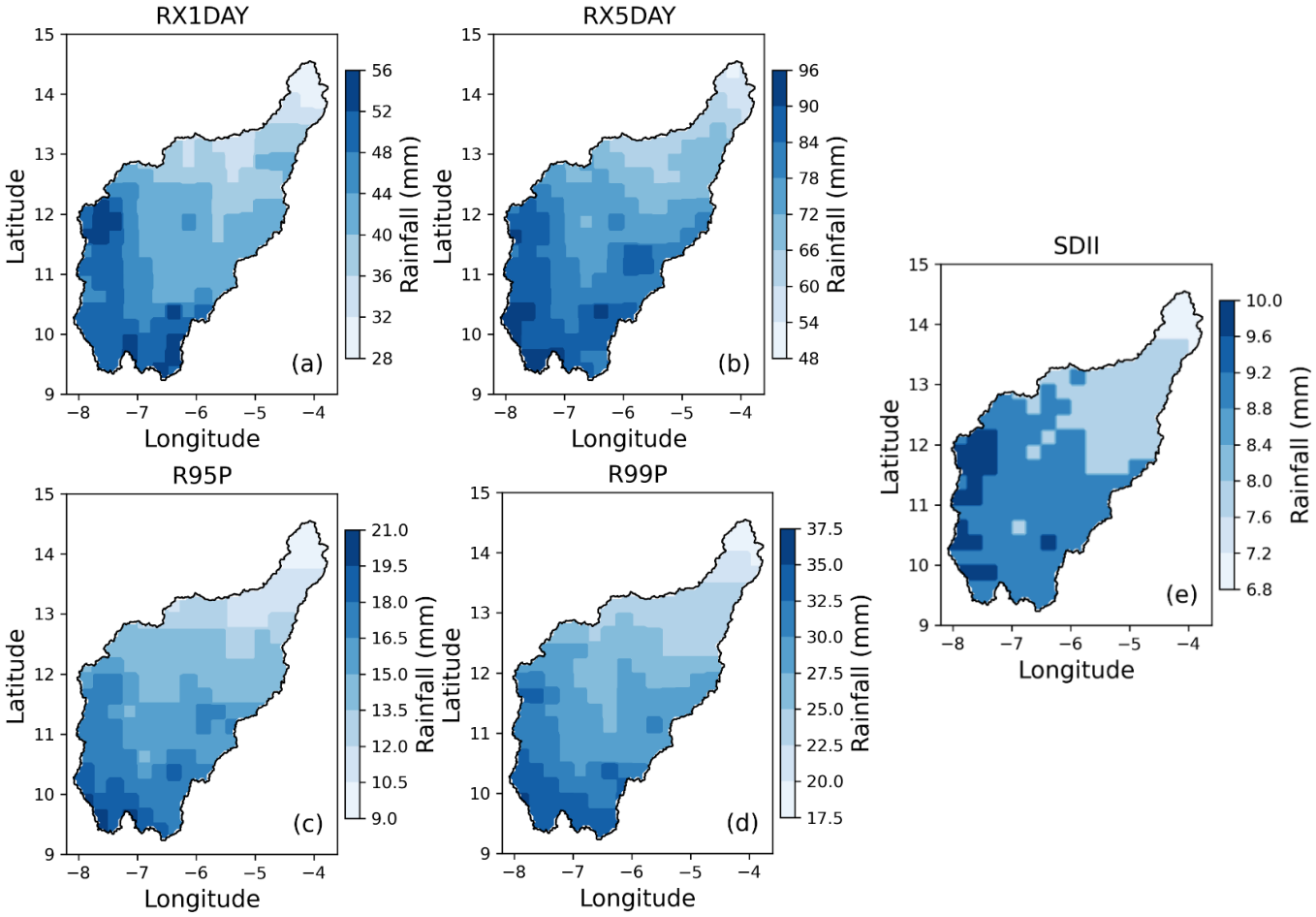

Figure 14: Spatial distribution of extreme rainfall indices across the study area (1991–2020): (a) maximum one-day precipitation, (b) maximum five-day precipitation, (c) very wet day, (d) extremely wet day, and (e) simple daily precipitation.

Such spatial disparities align with findings from across West Africa. For instance, ref. (Houngnibo et al., 2023c) reported increasing trends in extreme rainfall and the multi-day maxima in Benin, while ref. (Ogunrinde et al., 2021) highlighted a significant rise in the RX1DAY and cumulative rainfall in Nigeria. Similarly, ref. (Ndiaye et al., 2023) observed a marked amplification of extreme rainfall events, particularly the RX5DAY and R95P, in the Senegal River Basin, consistent with broader trends across southern West Africa.

4.6. Trend and significance of extreme indices

The spatial analysis of the Sen's slope trends in Figure 15 reveals notable variability across the study area in the extreme rainfall indices. The RX1DAY index in panel (a) highlights a positive

and significant trend of the maximum one-day rainfall in the northern and northeastern regions, indicated by black dots ($p < 0.05$). The RX5DAY index in panel (b) shows a significant upward trend in the five-day maximum precipitation in the northeastern and central regions, with a notable concentration of significant increases, while southern areas exhibit decreasing trends of lesser magnitude. Conversely, decreasing trends are observed in the southern part of the basin, with moderate significance. For the R99P index in panel (d), an increasing trend and significant rise in extremely wet days are concentrated in the northeastern areas. However, declining trends dominate the southwestern regions, indicating a reduction in extreme rainfall occurrences. The R95P index in panel (c) shows a positive and significant trend of very wet days in the northeastern part of the basin, while the southern regions exhibit moderate to negative trends. The SDII index in panel (e) displays an upward trend in the daily rainfall intensity in the central and northeastern zones, with significant increases near the basin's center. In contrast, the southern areas reveal moderate or no meaningful changes in the rainfall intensity. This variation suggests uneven rainfall intensification across the basin.

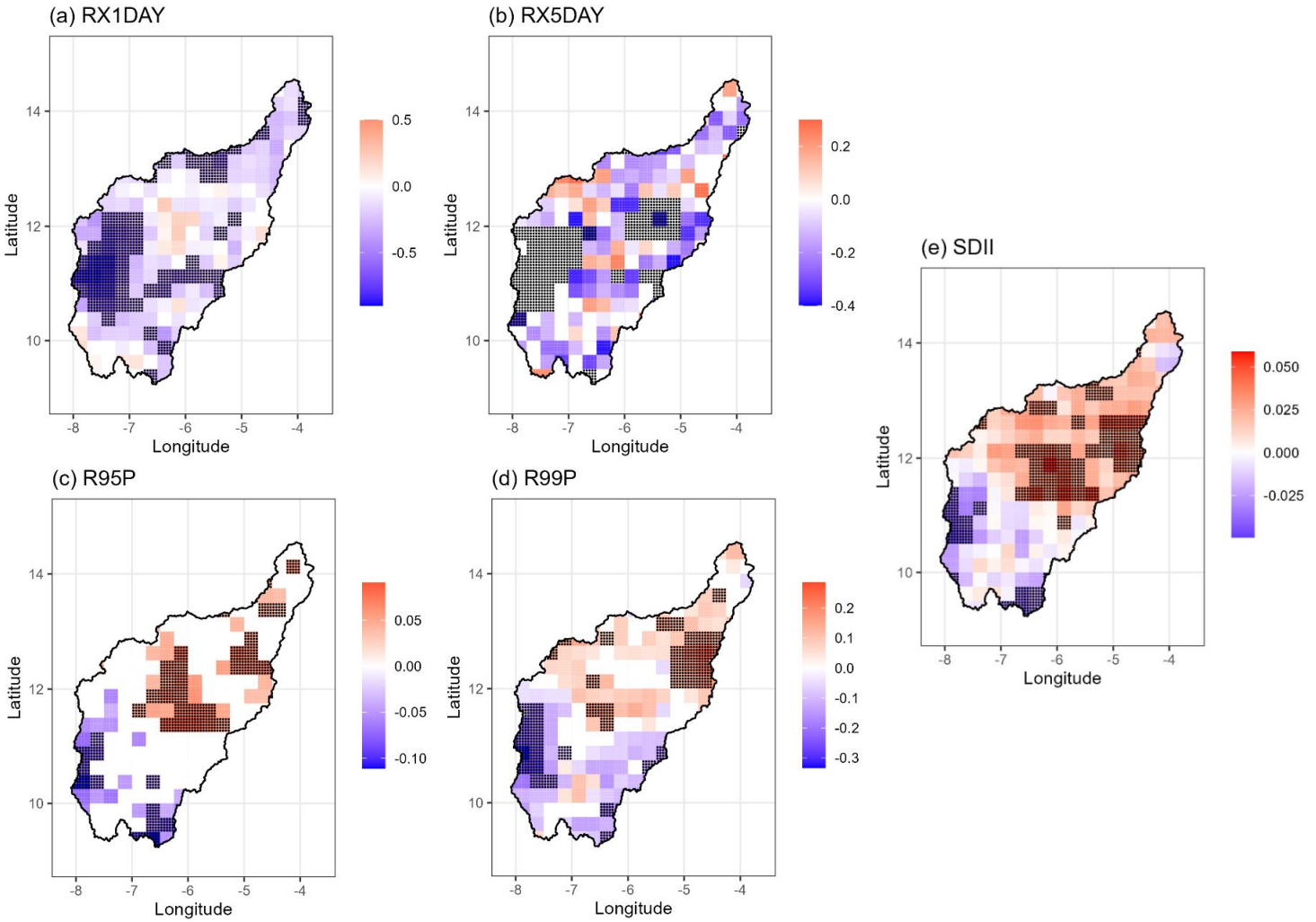

Figure 15: Sen's slope and trends for the BRB from 1991 to 2020: (a) RX1DAY, (b) RX5DAY, (c) R95P, (d) R99P, and (e) SDII. Black dots indicate statistically significant trends ($p < 0.05$), showing the spatial patterns of changes in extreme rainfall indices across.

4.7. Interannual variation and trends in extreme rainfall indices

The interannual trends of the extreme rainfall indices in the BRB from 1991 to 2020 demonstrate considerable variability over time and across regions (Figure 13). The RX5DAY index shows a significant decline (Sen's slope = −0.15 mm/year, $p < 0.01$), particularly after 2000, indicating a reduction in prolonged heavy rainfall events that are typically associated with flooding. This trend suggests potential implications for flood risk mitigation but also raises concerns about shifts in the rainfall distribution over shorter durations. Similarly, the RX1DAY index highlights a decline (Sen's slope = −0.2 mm/year, $p < 0.01$), particularly after 2012, reinforcing the reduction in extreme rainfall intensities in some regions. Conversely, the R99P index displays a significant upward trend (Sen's slope = 0.07 mm/year, $p < 0.001$), pointing to an increase in extreme rainfall events, especially in the northeastern parts of the basin. These results align with broader regional observations, such as those in (Quenum et al., 2021), which reported increasing rainfall extremes across West Africa. The SDII index shows a weak positive trend (Sen's slope = 0.008 mm/day/year), indicating only slight changes in the rainfall intensification. Meanwhile, the R95P index, representing very wet days, shows no significant trend (Table 9), suggesting that the frequency of moderately extreme rainfall events has remained stable.

Table 9: Modification of the Mann–Kendall test and Sen's slope estimation from 1991 to 2020.

Indices	***p*-Value**	**Zc**	**Sen's Slope (mm/Year)**	**Tau**	**Var(s)**	**Units**
R95P	11×10^{-2}	0.67	0	0.08	500.15	mm/year
R99P	17×10^{-5}	1.18	0.07	0.15	300.42	mm/year
SDII	43×10^{-3}	0.98	0.008	0.13	739.44	mm/year
RX1DAY	10^{-3}	−1.14	−0.2	−0.15	382.43	mm/year
RX5DAY	11×10^{-3}	−1.18	−0.15	−0.15	680.5	mm/year

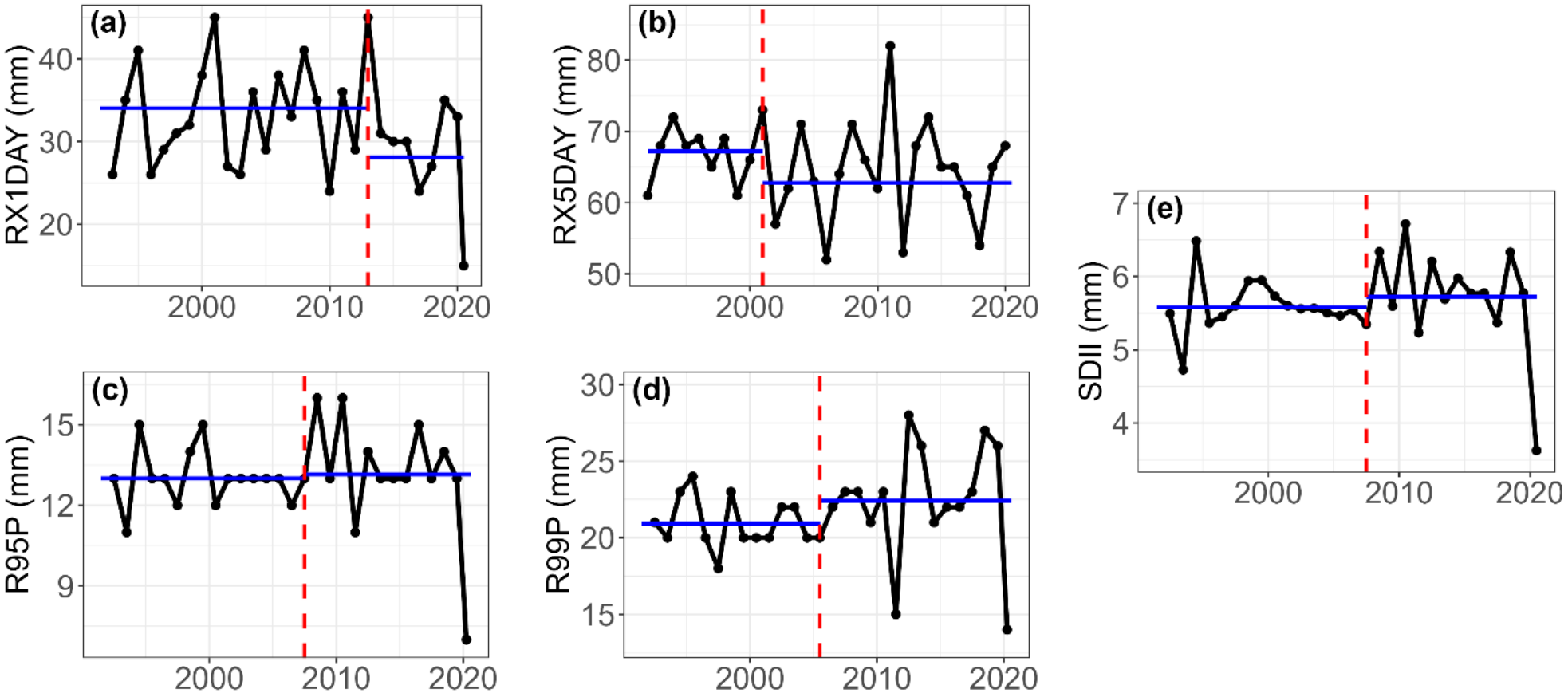

Figure 16: Breakpoint results for extreme precipitation over the BRB: (a) RX1DAY, (b) RX5DAY, (c) R95P, (d) R99P and (e) SDII.

The breakpoints in these indices further highlight structural changes in the rainfall patterns (Figure 16). The RX1DAY index shows a marked decline after 2012, while the R99P and SDII indices indicate shifts around 2005 and 2007, respectively (Table 10). These breakpoints reflect changes in the rainfall dynamics that may be linked to broader regional climatic shifts, as noted in (Chaibou Begou et al., 2016).

These findings illustrate the complex nature of hydroclimatic variability in the BRB. Southern regions are experiencing reduced extreme rainfall events, while the northeastern areas face increased risks of intense rainfall and flooding.

Table 10: Shifts in the extreme precipitation indices in the BRB during the period 1991–2020.

Index	p-Value	Breakpoint
R95p	1	2007
R99p	0.4229	2005
SDII	0.487	2007
RX1DAY	0.66	2012
RX5DAY	0.68	2000

4.8. Analysis of extreme flow characteristics

Figure 17 provides insight into the fluctuations in water availability across the BRB from 1991 to 2020. The SFI effectively highlights periods of hydrological surplus and deficit, which is critical for understanding the basin's response to climatic and hydrological changes. Between 1991 and 2000, the BRB experienced multiple years of below-average discharge, as indicated by the prevalence of negative SFI values. This period reflects prolonged hydrological deficits, which align with the region's well-documented droughts during the 1990s. The deficits observed are consistent with the findings in (Chaibou Begou et al., 2016), which reported significant reductions in river flows during this time due to declining rainfall. From 2001 onwards, there is a noticeable shift, with the more frequent positive SFI values reflecting above-average discharge. These findings align with broader trends in West Africa, where a partial recovery of rainfall has been observed since the late 1990s, as documented in (Ndiaye et al., 2023). The alternating periods of wet and dry years, especially the shifts between surplus and deficit conditions post-2010, underscore the basin's vulnerability to climatic variability.

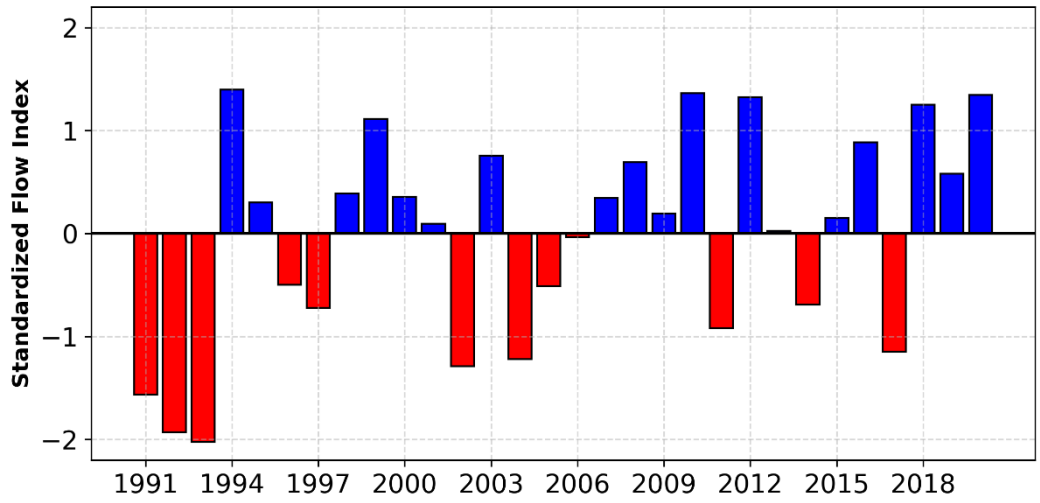

Figure 17: Standardized flow index over the BRB (1991–2020).

- Trends and interannual variability of extreme flows

The trends and interannual variability of extreme flows in the BRB from 1991 to 2020 reveal significant changes in high-flow conditions, as shown in Table 11. The analysis of indices such as the Q95P, Q99P and Qmax highlights upward trends across all three metrics, suggesting an intensification of extreme flow events over the study period. The Q95P index exhibits a significant positive trend (Sen's slope = 19.75 m^3/s/year, $p < 0.001$), indicating an increase in moderately extreme flows. Similarly, the Q99P index shows a significant rise (Sen's slope = 21.04 m^3/s/year, $p < 0.001$). This suggests a growing frequency of extreme flood conditions, which may pose heightened risks of flooding in the basin. The peak discharge, represented by the Qmax index, also demonstrates a significant upward trend (Sen's slope = 20.92 m^3/s/year, $p < 0.001$). This trend points to an intensification of the maximum river flows over time, consistent with the increasing variability of extreme flow conditions observed in the basin. The rising trends in all the indices underscore a pattern of hydrological intensification, characterized by increased flow extremes. When compared with historical discharge records and neighboring basins, these recent increases in the Q95P and Q99P are notably larger than past norms, highlighting a broader regional intensification of high-flow conditions (Intergovernmental Panel On Climate Change (IPCC), 2023a; Kundzewicz and Robson, 2004).

Table 11: Sen's slope and modified Mann–Kendall test on the extreme flow for the 1991 to 2020 period.

Indices	*p*-Value	Zc	Sen's Slope	Tau	Var(s)	Units
Q95P	24×10^{-9}	2.12	19.75	0.28	494.27	$m^3/s/year$
Q99P	68×10^{-10}	2.35	21.04	0.3	518	$m^3/s/year$
Qmax	25×10^{-10}	2.35	20.92	0.3	490	$m^3/s/year$

- Breakpoint Detection on the Trends of Extreme Flows

The analysis of the breakpoints for the extreme flow indices in the BRB identifies 1993 as an important year for significant changes in the hydrological regime. Both Pettitt's test and the SNHT consistently detect this shift across all the indices, as shown in Table 12. The Q95P and Q99P indices, which represent high-flow and very-high-flow days, exhibit statistically significant breakpoints ($p < 0.05$) in 1993. This indicates a noticeable increase in extreme flow events after this year, as reflected in figure 18. Similarly, the Qmax index, representing the peak discharge, also shows a breakpoint in 1993, signaling a transition toward higher maximum flows. These results suggest that 1993 marks a turning point in the basin's flow patterns, potentially due to climatic variability or environmental changes.

Table 12: Breakpoints related to the extreme flow indices (1991–2020).

Indices	Pettitt's Test		SNHT	
	p-Value	Breakpoint	p-Value	Breakpoint
Q95P	0.013	1993	0.013	1993
Q99P	0.042	1993	0.0166	1993
Qmax	0.042	1993	0.139	1993

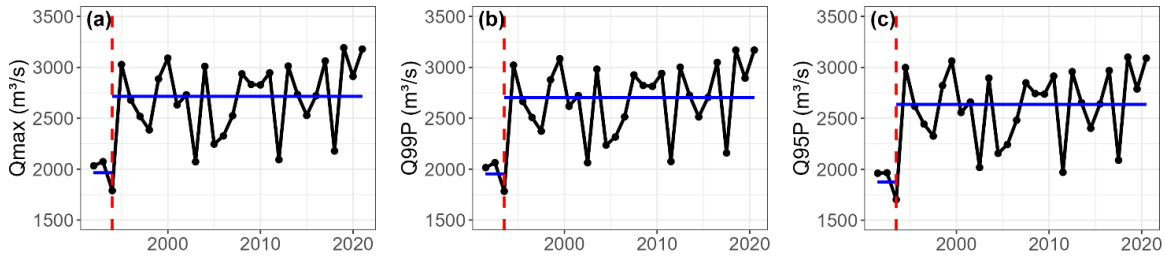

Figure 18 : Breakpoint results for the extreme flows over the BRB (1991–2020): (a) Qmax, (b) Q99P and (c) Q95P

4.9. Discussion

This study provides a detailed analysis of the extreme rainfall trends in the BRB from 1991 to 2020, highlighting the significant spatial and temporal variability in the rainfall indices. These findings align with broader regional and global patterns of hydroclimatic variability driven by climate change. For instance, the decline in the RX5DAY and RX1DAY indices, particularly after 2012, parallels trends observed in other semi-arid regions, such as the Sahel, where reduced prolonged rainfall events have impacted agricultural productivity and water availability (Nicholson, 2013). Conversely, the increase in the R99P and SDII, indicating intensified short-duration rainfall, mirrors findings from basins such as the Senegal River Basin and the Ganges, where extreme rainfall events have become more frequent due to shifting climatic regimes (Mishra and Singh, 2011; Sylla et al., 2015). The breakpoints identified (e.g., 2000 for the RX5DAY and 1993 for the extreme flows) further highlight structural changes in rainfall dynamics that may be linked to global climate drivers, such as the El Niño–Southern Oscillation (ENSO) (Ward et al., 2014).

The 1993 breakpoint in the extreme flows (Q95P, Q99P, and Qmax) may reflect broader climate regime shifts linked to large-scale ocean–atmosphere interactions, such as changes in Atlantic sea surface temperatures or ENSO cycles, which have been shown to alter precipitation patterns in West African river basins (Nicholson, 2013). From a management perspective, recognizing this hydrological turning point underscores the need for adaptive measures such as floodplain zoning and infrastructural investments to mitigate heightened flood risks and bolster water security in the BRB (Di Baldassarre et al., 2019).

These shifts present challenges for water resource management in the BRB. The rise in the R99P underscores the need for improved flood risk management in northeastern regions, including the construction of retention basins and enhanced floodplain zoning. Meanwhile, the declining RX5DAY trends in southern areas call for adaptive strategies to sustain rain-fed

agriculture, such as efficient irrigation systems and water conservation measures (Biasutti, 2019; Sultan and Gaetani, 2016). Methodologically, this study leveraged the modified Mann–Kendall test to ensure reliable trend detection, accounting for autocorrelation. However, limitations in the satellite-derived CHIRPS data highlight the need to integrate ground-based observations in future research to capture localized extremes more accurately (Dinku et al., 2010).

The results of this study provide a valuable framework for policymakers to develop climate adaptation strategies, balancing flood mitigation with water resource sustainability. Further cross-basin comparisons, particularly with the Senegal and Nile Basins, could enhance regional adaptation efforts.

4.10. Partial conclusion

The BRB has experienced significant hydroclimatic variability from 1991 to 2020, characterized by significant shifts in rainfall patterns and river discharge. Southern regions experienced reduced prolonged rainfall events, while northeastern regions faced intensified short-duration rainfall. These findings underscore the need for localized water resource management strategies tailored to these distinct regional patterns. Methodologically, the use of the modified Mann–Kendall test provided a robust trend analysis by addressing the autocorrelation in the time series data. This approach ensures reliable identification of shifts, such as the 1993 breakpoint for the river discharge and the 2000 breakpoint for the RX5DAY, which reflect broader climatic drivers like the El Niño–Southern Oscillation (ENSO). Comparing these trends to other semi-arid basins, such as the Senegal and Niger, highlights the BRB's heightened vulnerability to hydroclimatic extremes. In terms of the river discharge, indices such as the Q95P, Q99P and Qmax (peak discharge) all demonstrated significant positive trends, indicating an intensification of extreme flow events over the study period. These outcomes align with broader regional studies that project substantial decreases in potential water availability across major West African river basins due to climate change (Carr et al., 2022; Fofana et al., 2022; Sylla et al., 2018). Additionally, our research indicates that future water resource developments in the BRB will be significantly impacted by climate change, pointing out the need for adaptive management strategies.

The observed trends in the BRB highlight the necessity of adaptive water resource management strategies that can effectively address both the reduction in heavy rainfall events and the intensification of extreme flow events. Implementing such strategies is crucial to mitigate the

impacts of climate variability and ensure sustainable water availability for agriculture, fisheries, and livelihoods in the region.

Chapter 5: Assessment of machine learning models describing the recorded water levels in the BRB

This chapter presents the configuration of input and output data critical for our comprehensive analysis of environmental parameters influencing hydrological patterns from 1983 to 2020. We detail the integration of diverse datasets, including total evaporation and water levels from multiple sites such as Pankourou and Mopti, into our models to maintain the chronological integrity of the data through sequential splits. Additionally, we outline our method of cross-validation, which enhances the robustness and predictive reliability of the models used in this study. This configuration process is crucial for training and evaluating our models on different temporal segments, thereby capturing the complex dynamics of hydrological changes over time. The results of this chapter have been submitted to a peer-reviewed journal.

5.1. Evaluation of the ML models for predicting water level fluctuations

5.1.1. Bayesian Optimization for Hyperparameter Tuning

Bayesian optimization has proven to be a highly efficient method for hyperparameter tuning in machine learning models, offering an intelligent approach for exploring and exploiting the hyperparameter search space. Unlike traditional methods such as grid search and random search, which can be computationally expensive and inefficient, Bayesian optimization builds a surrogate probabilistic model (often a Gaussian Process) to guide the selection of hyperparameters, progressively refining the model based on past performance (Snoek et al., 2012). This process allows the method to focus on regions of the hyperparameter space that are more likely to contain the optimal settings, thus balancing exploration and exploitation.

In this chapter Bayesian optimization was applied to RF, GB, SVR and XGB. The hyperparameters for these models, such as the number of estimators (n_estimators), learning rate, maximum tree depth (max_depth), and subsampling rate (subsample), were fine-tuned using this method, as outlined in Table 13. Notably, Bayesian optimization handled complex hyperparameters for non-tree-based models like SVR, where parameters C, epsilon, and gamma were adjusted to optimize the margin between predicted and actual values. In Table 14, the optimal hyperparameters varied significantly between Fold 1 and Fold 2, reflecting the model's adaptability to different training data characteristics. For example, RF in Fold 1 achieved its best performance with a shallower tree (max_depth = 10), while in Fold 2, a deeper tree (max_depth = 36) was selected. Similarly, XGB in Fold 1 selected a learning rate of 0.0597 and n_estimators of 304, while Fold 2 required a smaller learning rate (0.01) but more estimators

(n_estimators = 446) for better generalization. These variations emphasize the importance of fold-specific tuning in time-series forecasting models, as the data distribution and complexity vary across different partitions.

The success of Bayesian optimization in this context demonstrates its value in improving both the accuracy and efficiency of ML models by focusing computational resources on the most promising areas of the hyperparameter space (Shahriari et al., 2016). This method's ability to dynamically adjust based on performance evaluations ensures that the selected models are both accurate and generalizable.

Table 13: Hyperparameter search space (Fold 1 and Fold 2).

Hyperparameter	RF (Fold 1)	RF (Fold 2)	GB (Fold 1)	GB (Fold 2)	SVM (Fold 1)	SVM (Fold 2)	XGB (Fold 1)	XGB (Fold 2)
n_estimators	100–500	100–500	100–500	100–500	N/A	N/A	100–500	100–500
max_depth	10–50	10–50	3–10	3–10	N/A	N/A	3–10	3–10
min_samples_split	2–10	2–10	N/A	N/A	N/A	N/A	N/A	N/A
min_samples_leaf	1–4	1–4	N/A	N/A	N/A	N/A	N/A	N/A
learning_rate	N/A	N/A	0.01–0.3 (log-uniform)	0.01–0.3 (log-uniform)	N/A	N/A	0.01–0.3 (log-uniform)	0.01–0.3 (log-uniform)
subsample	N/A	N/A	0.6–1.0	0.6–1.0	N/A	N/A	0.6–1.0	0.6–1.0
colsample_bytree	N/A	N/A	N/A	N/A	N/A	N/A	0.6–1.0	0.6–1.0
C	N/A	N/A	N/A	N/A	0.1–1000 (log-uniform)	0.1–1000 (log-uniform)	N/A	N/A
gamma	N/A	N/A	N/A	N/A	0.0001–1.0 (log-uniform)	0.0001–1.0 (log-uniform)	N/A	N/A
epsilon	N/A	N/A	N/A	N/A	0.01–0.1	0.01–0.1	N/A	N/A

Table 14: Best Hyperparameters selected (Fold 1 and Fold 2) using Bayesian optimization method.

Model	Best Hyperparameters (Fold 1)	Best Hyperparameters (Fold 2)
RF	max_depth: 10, min_samples_leaf: 1, min_samples_split: 2, n_estimators: 292	max_depth: 36, min_samples_leaf: 3, min_samples_split: 7, n_estimators: 307
GB	learning_rate: 0.0487, max_depth: 3, n_estimators: 203, subsample: 1.0	learning_rate: 0.3, max_depth: 3, n_estimators: 500, subsample: 0.6
SVR	C: 0.1, epsilon: 0.1, gamma: 0.15	C: 33.2, epsilon: 0.01, gamma: 0.005
XGB	colsample_bytree: 0.6, learning_rate: 0.06, max_depth: 10, n_estimators: 304, subsample: 0.6	colsample_bytree: 0.6, learning_rate: 0.01, max_depth: 6, n_estimators: 446, subsample: 0.65

5.1.2. Cross-Validation Performance of ML Models for Fold 1

Figure 19 provides a comprehensive evaluation of four ML models, RF, GB, SVM, and XGB applied to predict water level in Mopti. In subfigures (a), (b), (c), and (d), the predicted versus observed water levels for the test set demonstrate that SVM (c) exhibits the highest predictive accuracy, with an R^2 value of 0.75, significantly outperforming RF (a) and the other models. Similarly, for the validation set, subfigures (e) through (h) reveal SVM (g) maintaining superior performance with an R^2 of 0.71, while RF (e) struggles with lower accuracy, as indicated by an R^2 of 0.4. The accompanying bar charts (i) through (l) highlight key model performance metrics. SVM consistently achieves the lowest RMSE and MAE, signifying its superior predictive capacity. The MAPE values, which indicate model bias, further confirm SVM's robustness. In terms of model efficiency, as reflected by the NSE metric (l), SVM again outperforms the alternatives, particularly in both test and validation sets. Table 15 present the numerical results of different statistical metrics use across the folds.

The feature importance analysis in subfigures (m), (n), and (o) reveals "WL_Sof" as the most important variable across all models, followed by "WL_Bou" and "Tot_evap," underscoring the consistency of feature contributions. Notably, SVM does not provide feature importance in the

same manner as tree-based models, as it relies on kernel functions rather than decision trees to capture complex relationships between features. Despite this, SVM's superior performance, as evidenced by its high R^2, lower error metrics, and model efficiency, establishes it as the most reliable and accurate model for water level prediction in this context.

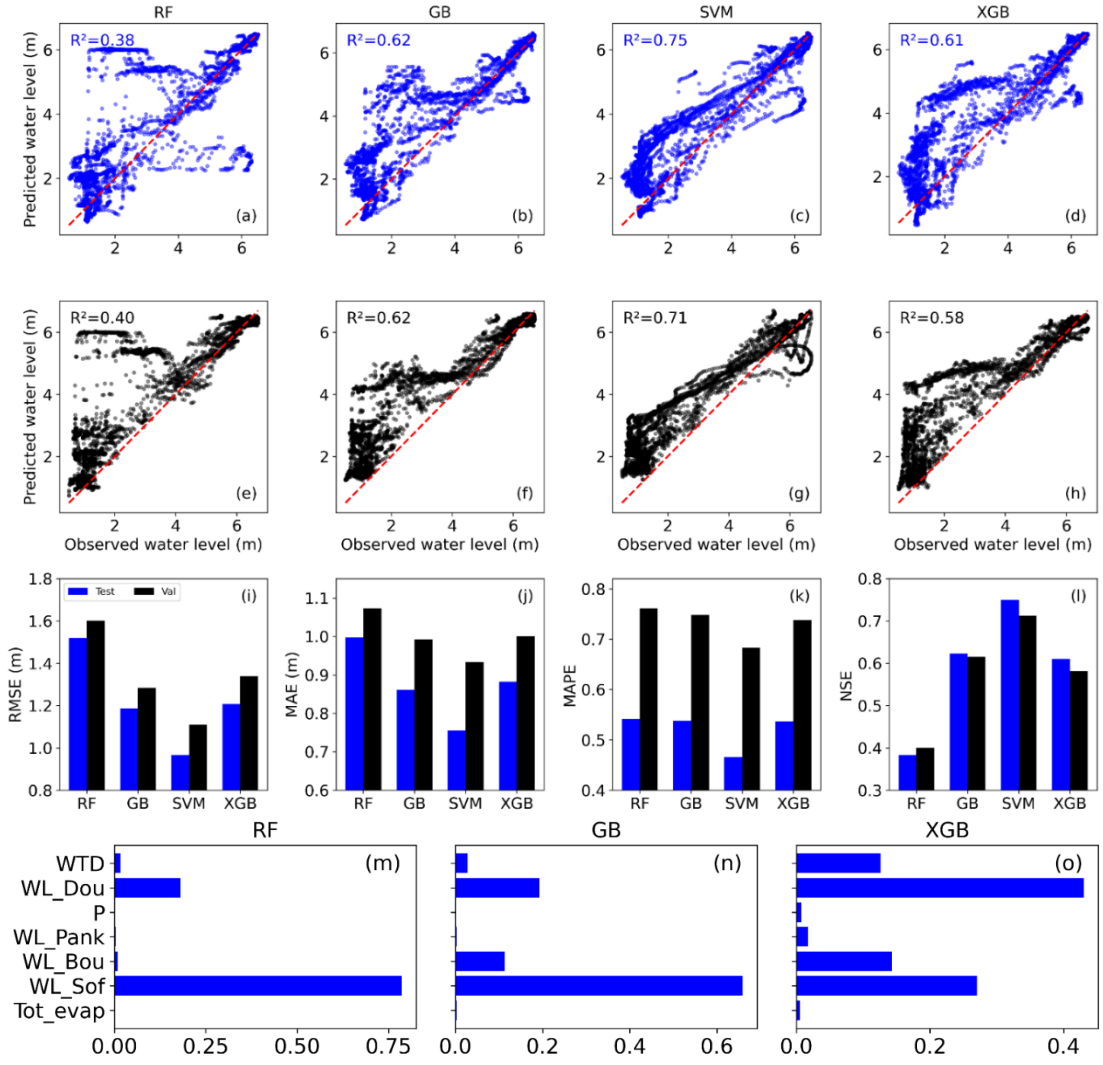

Figure 19: Evaluation of ML models for predicting water levels – Fold 1; (a), (b), (c) and (d): predicted vs. observed water levels using RF, GB, SVM and XGB algorithms, respectively for testing set; (e), (f), (g) and (h): predicted vs. observed water levels using RF, GB, SVM and XGB algorithms, respectively for validation set; (i) RMSE values for each model; (j) MAE values demonstrating model fit; (k) MAPE values showing model bias; (l) NSE values indicating model efficiency; feature importance in (m) RF; (n) GB; and (o) XGB.

Table 15: Evaluation results for the test set and validation set - Fold 1.

	Test				Validation			
Model	RMSE	MAE	MAPE	NSE	RMSE	MAE	MAPE	NSE
RF	1.52	1	0.54	0.38	1.6	1.07	0.76	0.4
GB	1.19	0.86	0.54	0.62	1.28	0.99	0.75	0.62
SVM	0.97	0.75	0.47	0.75	1.11	0.93	0.68	0.71
XGB	1.21	0.88	0.54	0.61	1.34	1	0.74	0.58

Figure 20 provides a comparative analysis of four ML models (RF, GB, SVM, and XGB) in predicting water levels for Mopti. SVM consistently outperforms other models across both test and validation sets, closely capturing seasonal and annual fluctuations, with minimal error metrics and high R^2 values. The findings underline SVM's adaptability to hydrological conditions and its potential for accurate water level predictions in similar contexts.

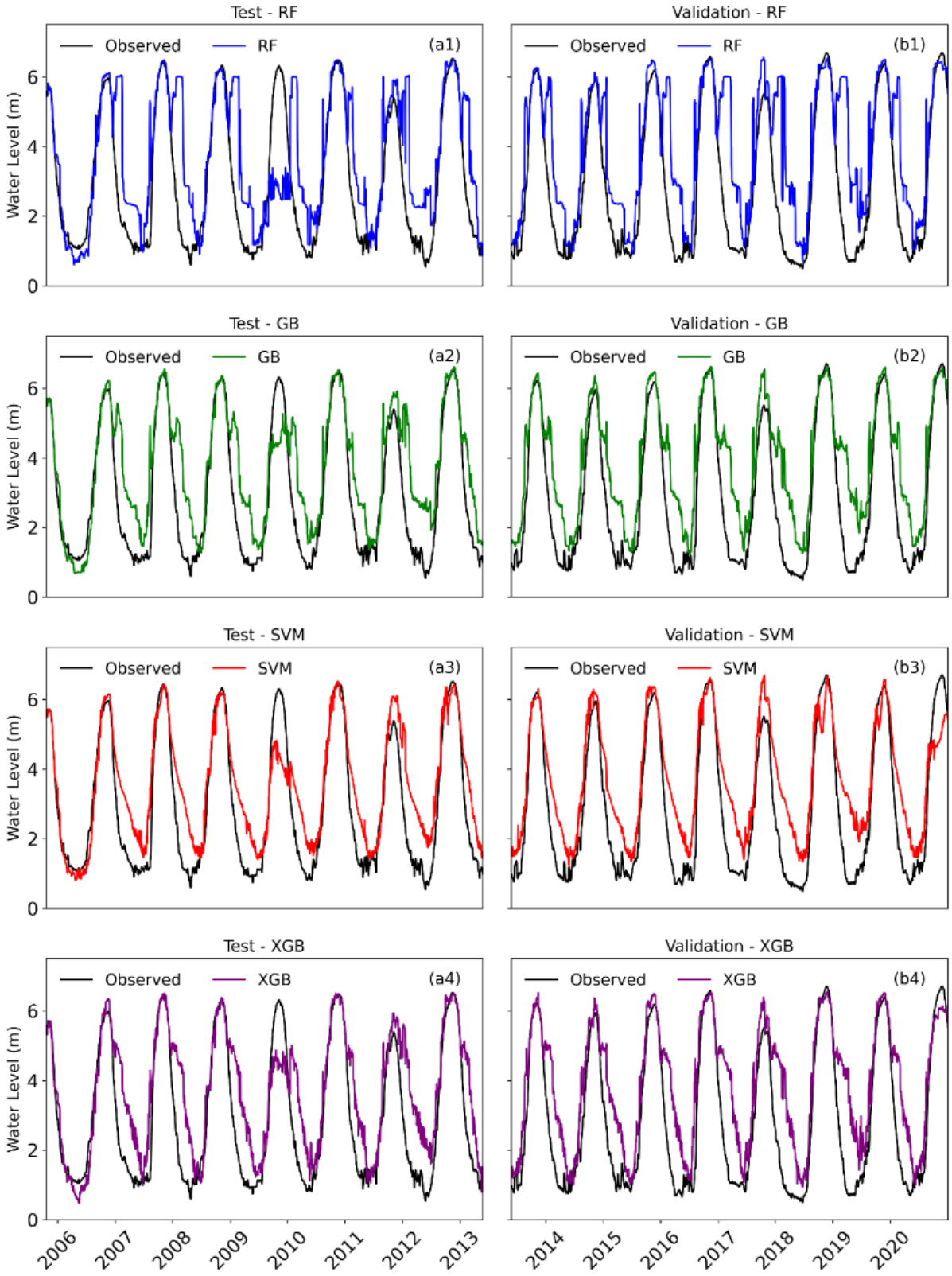

Figure 20: Comparative performance of water level predictions (2013-2020) – Fold 1; (a1), (a2), (a3) and (a4): model predictions vs. observed water levels using RF, GB, SVM and XGB algorithms, respectively for testing set; (b1), (b2), (b3) and (b4): model predictions vs. observed water levels using RF, GB, SVM and XGB algorithms, respectively for validation set.

5.1.3. Cross-validation performance of ML models for fold 2

Figure 21 evaluates the performance of four different ML models used in this investigation for predicting water level at the outlet of the BRB. We used seven predictors including water table depth (WTD), total evaporation (Tot_evap), CHIRPS rainfall (P), and water level (WL) at four upstream locations in the basin (Pankourou, Bougouni, Douna, Sofara) to train our model and predict the water level at Mopti. In the previous section, it is detailed that Bayesian optimization was utilized to efficiently tune hyperparameters for ML models, focusing computational resources on the most promising areas of the hyperparameter space and adapting to different training data characteristics across folds. To enhance the models' robustness and predictive reliability, cross-validation was employed by splitting the dataset sequentially while maintaining its chronological integrity. Specifically, in the Fold 1, the first 60% of the data was allocated for training, the next 20% for testing, and the final 20% for validation. In the Fold 2, the first 20% was used for testing, the middle 60% for training, and the last 20% for validation. This approach allowed the models to be trained and evaluated on different temporal segments, improving their ability to model complex hydrological patterns. The scatter plots in Figure 17a-h compare the predicted water levels from each model with observed water levels for testing and validation set. The R^2 value for each case is presented on the plot with their high values suggesting a very high level of accuracy in predictions and the right set of parameters chosen as the predictors. For Fold 1 the results are presented in the Table 15, SVM model achieved superior predictive accuracy (NSE > 0.7) for water level predictions, outperforming other models with the highest R^2 values and lowest error metrics in both test and validation sets. For Fold 2, all models demonstrated high performance (Figure 21), with NSE values exceeding 0.7, indicating strong predictive accuracy in estimating water levels. On the test set, the SVM model achieved the lowest RMSE and MAPE (Figure 21i-k), while all models attained high NSE values ranging from 0.91 to 0.94. Similarly, on the validation set, the RF model achieved the highest NSE and the lowest RMSE (Figure 21e). Despite slight variations in the performance metrics such as GB having a higher RMSE and MAPE, all models performed well overall. Figures 21m-o present the feature importance for RF, GB, and XGB models in Fold 2, respectively. The charts highlight the WL_Bou as a critical predictor across all models, particularly in RF and XGB. WL_Sof, along with Tot_evap and WTD, also played significant roles in influencing the predictive outcomes. Although RF and XGB models demonstrated satisfactory predictive performance in this study, it is important to interpret their superiority with caution. These models were not directly compared to traditional statistical approaches applied to the same dataset and study periods. Therefore, the results should be considered as

relative indications rather than absolute evidence of superiority. Nevertheless, recent literature in hydrology confirms that machine learning methods generally outperform classical statistical models in capturing the nonlinear dynamics of hydroclimatic processes (Hassani et al., 2024; Yaseen et al., 2019). This study reinforces that trend while also highlighting the need for complementary investigations based on multi-method comparisons.

The feature importance in Fold 1 similarly emphasized the significance of WL_Bou and WL_Sof as key predictors, consistently impacting model performance across different folds. This finding suggests that these variables are essential for accurate water level predictions, regardless of temporal splits, reinforcing their importance in managing hydrological patterns within the BRB.

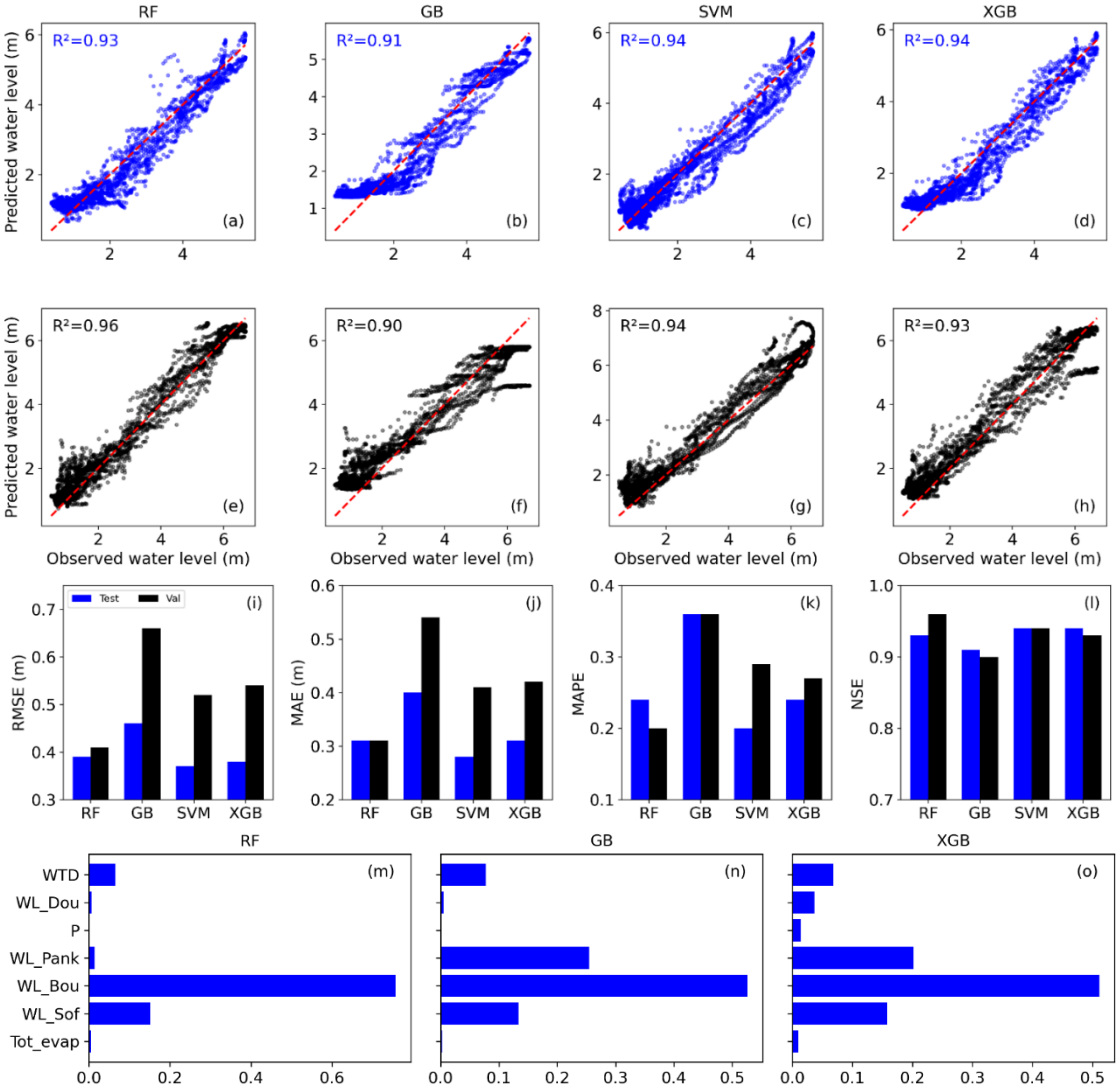

Figure 21: Evaluation of ML models for predicting water levels – Fold 2; (a), (b), (c) and (d): predicted vs. observed water levels using RF, GB, SVM and XGB algorithms, respectively for testing set; (e), (f), (g) and (h): predicted vs. observed water levels using RF, GB, SVM and XGB algorithms, respectively for validation set; (i) RMSE values for each model; (j) MAE values demonstrating model fit; (j) NSE values indicating model efficiency; (k) MAPE values showing model bias; feature importance in (m) RF; (n) GB; and (o) XGB.

Figure 22 shows that all models performed well, capturing both seasonal variations and annual fluctuations in Mopti. During the implementation of Fold 1, only the SVM model exhibited strong performance, closely matching the observed water levels with high accuracy. This comparative analysis not only confirms the robustness of the developed SVM model but also demonstrates the model's adaptability to hydrological conditions similar to those in the BRB. It highlights the importance of selecting appropriate features and tuning model parameters to optimize prediction accuracy for water level management. Given the successful application of these models in the BRB, they hold potential for adaptation to similar hydrological settings. However, further validation under varying conditions is essential to ensure their broader applicability, particularly in regions with comparable environmental and climatic dynamics.

Note that these findings are based on the specific setup and input data used in this study. As shown in Figure 21m-o, the importance of features varies for each model, indicating that the performance of these algorithms can be highly sensitive to the choice and configuration of input variables. Therefore, while the results indicate the superiority of all the models in this context, this conclusion may not be generalizable across all scenarios. The performance observed here should be validated under different conditions and local datasets to confirm its broader applicability.

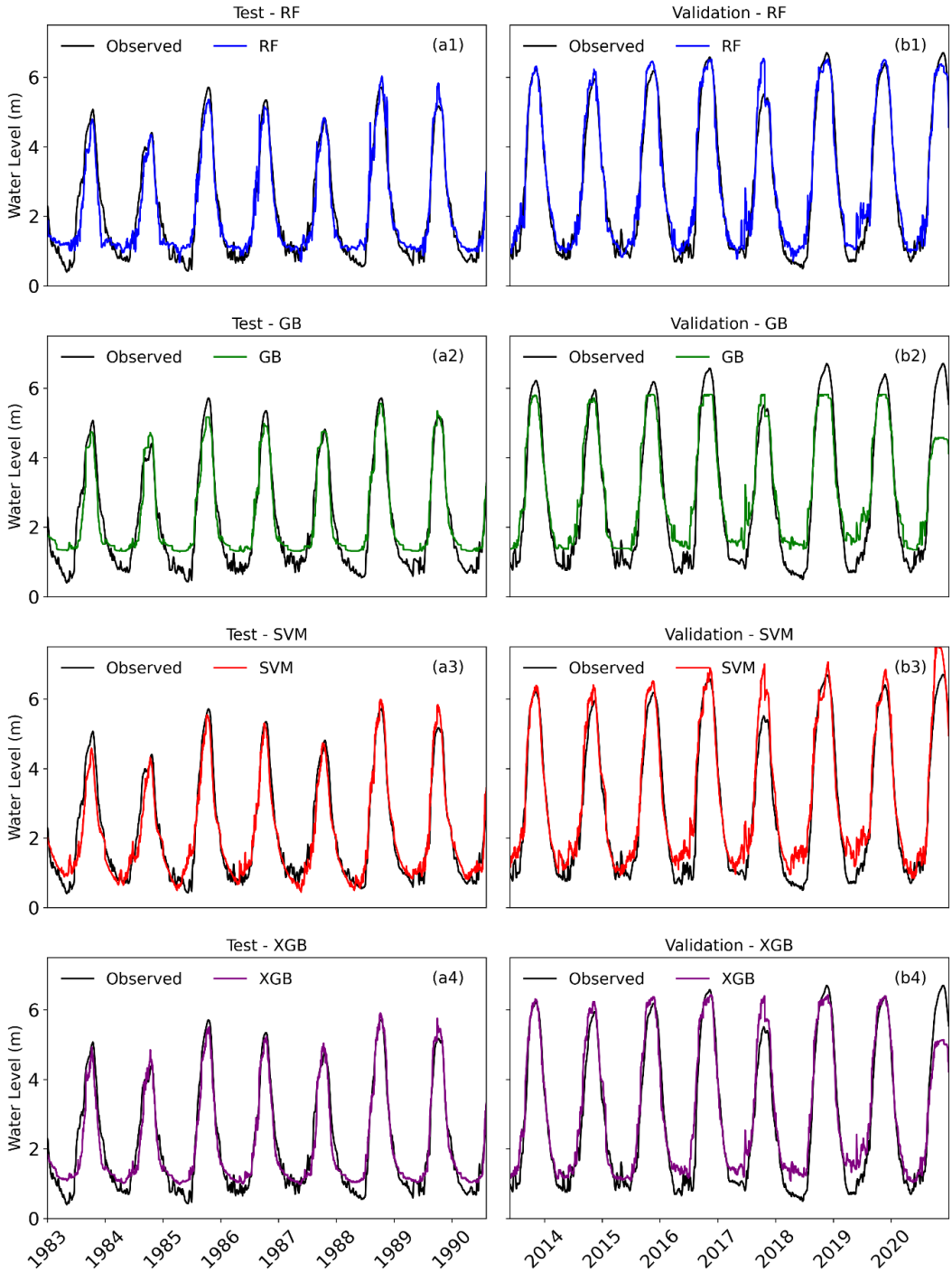

Figure 22: Comparative performance of water level predictions – Fold 2; (a1), (a2), (a3) and (a4): model predictions vs. observed water levels using RF, GB, SVM and XGB algorithms, respectively for testing set; (b1), (b2), (b3) and (b4): model predictions vs. observed water levels using RF, GB, SVM and XGB algorithms, respectively for validation set.

5.2. Correlation Heatmap

In the correlation analysis between the input features and the water level at Mopti (WL_Mop), several key relationships emerge, providing insight into the factors influencing this target variable (Figure 23). The most positively correlated features with

WL_Mop include WL_Bou, WL_Sof, and WL_Dou. Among these, WL_Bou demonstrates the highest positive correlation, suggesting that as water levels at Bougouni increase, water levels at Mopti follow a similar trend. This is reasonable since these locations are part of the same river system, and upstream water levels naturally affect downstream locations. Similarly, WL_Sof and WL_Dou show significant positive correlations, reinforcing the idea that water levels at these nearby stations are interrelated and indicative of broader hydrological conditions in the region. On the other hand, Tot_evap shows a strong negative correlation with WL_Mop. This inverse relationship indicates that higher evaporation rates contribute to lower water levels at Mopti. This makes intuitive sense, as evaporation decreases water availability, leading to decreased water levels. Additionally, WTD also shows a weaker, but still notable negative correlation, indicating that higher water table depths (deeper groundwater levels) are associated with reduced surface water levels at Mopti. The correlation analysis reveals that upstream water levels at Bougouni, Sofara, and Douna are reliable indicators of water levels at Mopti. Conversely, increased evaporation and deeper water tables are associated with lower water levels. Understanding these relationships is important for hydrological modeling and water resource management in the region, providing a foundation for predicting water level fluctuations based on upstream conditions and environmental factors.

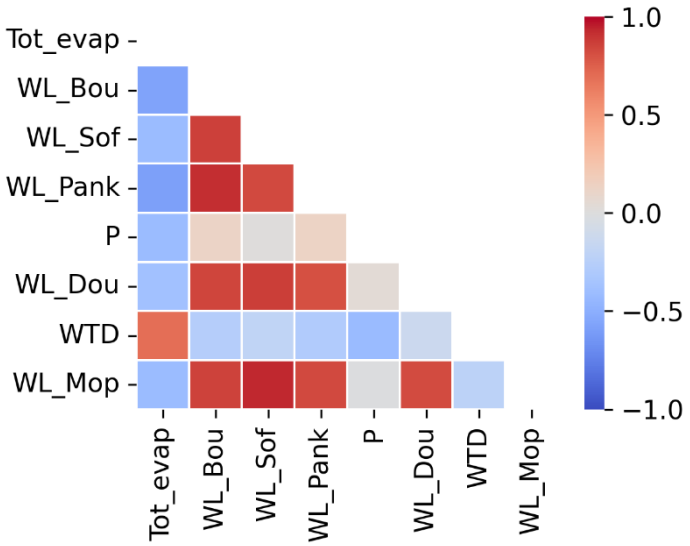

Figure 23: Correlation heatmap between water level at Mopti and other features.

5.3. Partial conclusion

This specific objective aimed to enhance hydrological monitoring and predictive capabilities within the Bani catchment, a transboundary region characterized by its complexity due to both natural and anthropogenic factors. Invoking satellite-based precipitation data of CHIRPS, we developed ML models to predict water level fluctuations at Mopti, the outlet of the BRB. These models were selected based on their robustness in handling non-linear and complex data structures, which is typical of hydrological datasets. All the ML models effectively captured water level dynamics and improved prediction abilities by averaging multiple decision trees which reduces the risk of overfitting (Xiang et al., 2023). SVM, particularly, may be the preferred model if the data variability remains challenging, but tuning the other models can also help achieve stable performance across all folds as shown by Saxena et al. (2023).

Our approach provides a framework that can be adapted to other regions facing similar socio-political challenges. As climate change and human impacts continue to complicate water management, the findings of this study contribute to a growing body of knowledge that supports the adoption of advanced technologies for sustainable hydrological practices (Shokri et al., 2023).

Chapter 6: Machine Learning and climate projections for assessing hydroclimatic extremes in the BRB

The BRB in southeastern Mali forms a vital tributary of the Niger River, supporting agricultural production, fisheries, and wetland ecosystems crucial to local livelihoods (Intergovernmental Panel On Climate Change (IPCC), 2023a; Traore, 2015). Its predominantly rain-fed agricultural system and reliance on seasonally fluctuating water resources underscore its importance in ensuring food security for surrounding communities (Ba, 2023; Bodian et al., 2020, 2018; Tefera et al., 2025). Because water availability directly affects planting cycles, fishing yields, and biodiversity in wetland habitats, maintaining a stable hydrological balance has long been pivotal to the region's socio-economic well-being (Nicholson, 2013; Okoronkwo et al., 2024).

Despite the BRB's significance, recent studies highlight mounting hydroclimatic stressors such as rising temperatures, erratic rainfall patterns, and soaring evaporation rates (Addor et al., 2020; Taylor et al., 2012). These factors contribute to water shortages during the critical agricultural season, while also increasing the risk of floods in low-lying areas 999. In combination with limited infrastructural resilience exemplified by vulnerable irrigation networks and insufficient flood-control systems these changing conditions threaten both food supply and economic stability (Demirel et al., 2009). Moreover, climate-induced alterations in water flow disrupt fish migration routes and degrade wetland ecosystems, jeopardizing local fisheries (Mahe et al., 2013; Roudier and Mahe, 2010b).

Socio-political uncertainties in Mali further exacerbate the challenges posed by climate change in the BRB (Ba, 2023; Kohnert, 2023). Persistent security concerns often impede systematic data collection, weakening the foundation of hydrological forecasting and long-term research in the basin. Infrastructure projects, including dams and irrigation schemes, continue to reshape flow regimes in ways that remain poorly understood, particularly when overlaid with uneven climate signals. The convergence of these pressures climatic, infrastructural, and socio-political raises the specter of severe water crises, with implications for livelihoods and regional stability.

Traditional statistical and deterministic models have been instrumental in evaluating historical hydrological trends, offering insights into seasonal flow variability and long-term water balances (Hamed and Ramachandra Rao, 1998). However, these models often struggle to capture non-linear dynamics introduced by abrupt climate changes, irregular data quality, and anthropogenic alterations to flow regimes (Kasei et al., 2010). Their reduced effectiveness under incomplete or inconsistent datasets further constrains their utility in conflict-prone

regions, leaving a gap in accurate, near-future forecasting for stakeholders who must make rapid decisions to adapt.

Machine learning (ML) methods, particularly RF and GB, have recently gained momentum in hydrological science for their capacity to model complex, multivariate processes and handle missing or inconsistent data. Studies across various African river basins show that ML approaches can enhance flood and drought predictions, optimize reservoir operations, and inform water allocation strategies (Banadkooki et al., 2019; Xiang et al., 2023). These methods also integrate seamlessly with climate projections, such as those from the CMIP6, to assess future hydroclimatic scenarios. However, little work has combined ML based water-level forecasting with abrupt change detection tools in the specific context of the BRB (Pettitt, 1979). Motivated by these pressing gaps, this study employs RF and GB to predict near-future water levels in the BRB under two CMIP6 emission scenarios SSP2-4.5 and SSP5-8.5. We introduce a water-level maximum (WLmax) index to identify potential flood-triggering events, and we apply the Pettitt test and modified Mann-Kendall procedure to detect abrupt and gradual hydrological shifts, respectively. Additionally, we investigate changes in three key precipitation indices to evaluate extended dry spells and intense rainfall events. By integrating advanced ML tools with cutting-edge climate scenario analyses, our study aims to provide timely, practical forecasts that can guide adaptation measures, optimize infrastructure planning, and reinforce community resilience.

Ultimately, this work seeks to offer a holistic perspective on the BRB's near-future hydrological regime, bridging data and methodological gaps to support more effective water resource management. Our findings are designed to inform policy-makers, local stakeholders and development agencies as they devise strategies to secure food production, safeguard biodiversity in wetland regions, and mitigate socio-economic vulnerabilities.

This chapter presents the development of ML models and the configuration of input and output data critical for our comprehensive analysis of environmental parameters influencing hydrological patterns from 1983 to 2020. We detail the integration of diverse datasets, including total evaporation and water levels from multiple sites such as Pankourou and Mopti, into our models to maintain the chronological integrity of the data through sequential splits. Additionally, we outline our method of cross-validation, which enhances the robustness and predictive reliability of the models used in this study. This configuration process is crucial for training and evaluating our models on different temporal segments, thereby capturing the complex dynamics of hydrological changes over time.

6.1. Evaluation of climate models

The Taylor diagram evaluates the performance of CMIP6 models in replicating rainfall patterns in the BRB using CHIRPS data as the reference (Figure 24). It assesses correlation coefficient, standard deviation, and root mean square error (RMSE). Most models demonstrate strong correlations above 0.8, indicating their ability to capture temporal rainfall variability. The ensemble mean (Ensmean) achieves the highest correlation (0.9), underscoring the advantage of combining models to reduce bias and improve reliability. In terms of standard deviation, the reference dataset shows a value of approximately 90. Models such as MPI-ESM1-2-HR and CanESM5 closely align with this, effectively representing rainfall variability. Conversely, ACCESS-CM2 and BCC-CSM2-MR deviate significantly, suggesting challenges in capturing regional-scale processes. The RMSE analysis confirms that the ensemble mean is closest to the reference, highlighting its overall accuracy. This aligns with previous findings by Sylla et al. (2018), demonstrating the robustness of ensemble approaches.

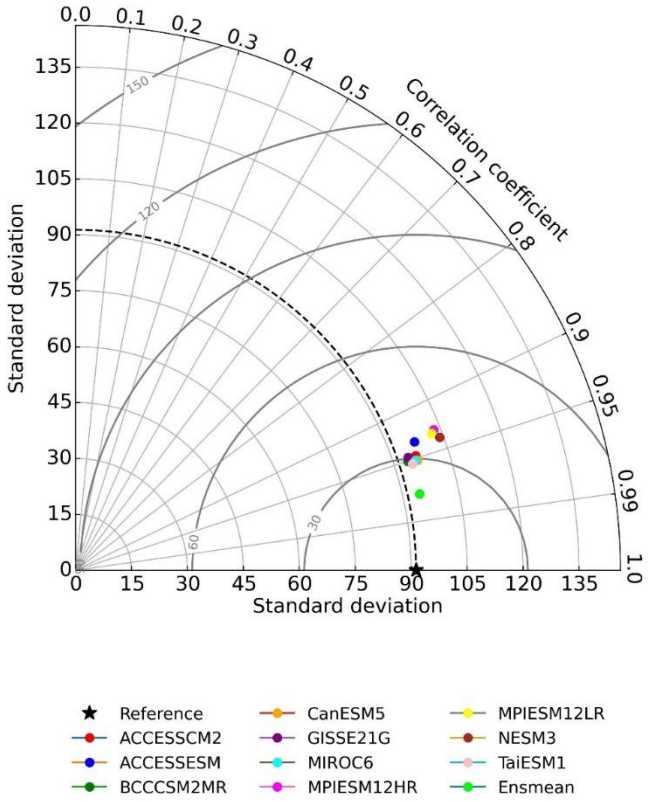

Figure 24: Taylor diagram comparing CMIP6 model performance against reference data.

6.2. Predictive model performance and validation

The validation results in Table 16 indicate that both RF and GB models perform satisfactorily in simulating water levels. With an NSE value of 0.76, RF slightly outperforms GB (NSE = 0.72), showcasing better accuracy in replicating observed water levels. These results underscore the importance of hyperparameter optimization, as the selected parameters for each model effectively capture the dynamics of the system.

Table 16: Model hyperparameters and NSE values.

Model	Hyperparameters	NSE (Validation)
RF	max_depth: 10, min_samples_leaf: 1, min_samples_split: 2, n_estimators: 100	0.76
GB	learning_rate: 0.02, max_depth: 4, n_estimators: 269, subsample: 0.97	0.72

Figure 25 validates the models against observed data, where RF and GB closely follow the seasonal trends. The minor deviations between the simulated and observed data suggest room for improvement, particularly in extreme events or transitions. RF's ability to maintain higher fidelity in such scenarios likely explains its superior NSE score.

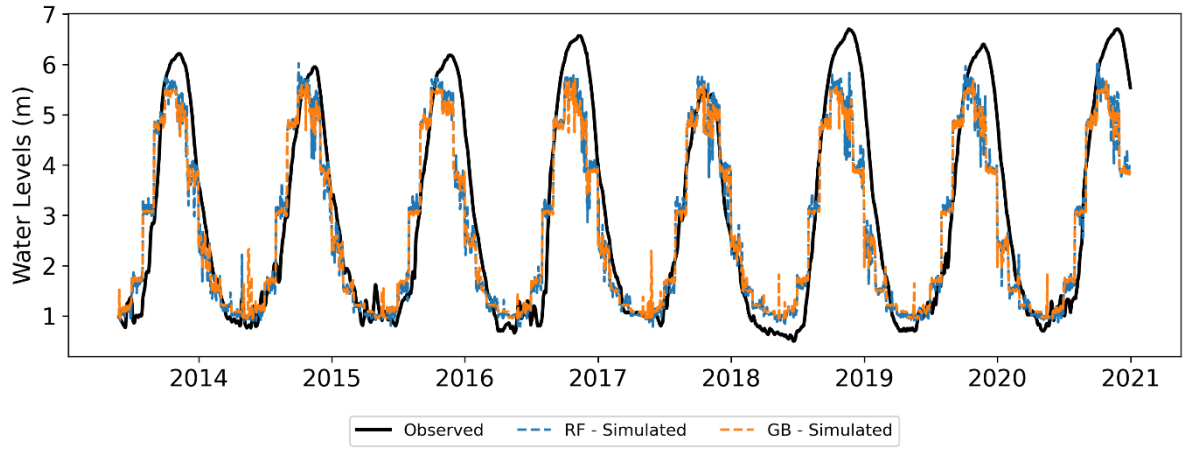

Figure 25: Simulated water levels compared to validation water levels using RF and GB models.

Figure 26 illustrates future water level projections under two scenarios. Both models predict consistent seasonal patterns, with RF and GB closely aligned, indicating their robustness for long-term simulation. However, RF slightly outperforms GB in capturing detailed variability, reinforcing its suitability for predictive tasks. These results highlight the reliability of ensemble

methods in hydrological modeling. RF's performance affirms its capacity to generalize well with minimal overfitting, while GB offers competitive accuracy with nuanced predictions. Such models are invaluable for water resource planning, particularly in a changing climate context, as suggested by previous studies (Pham et al., 2021).

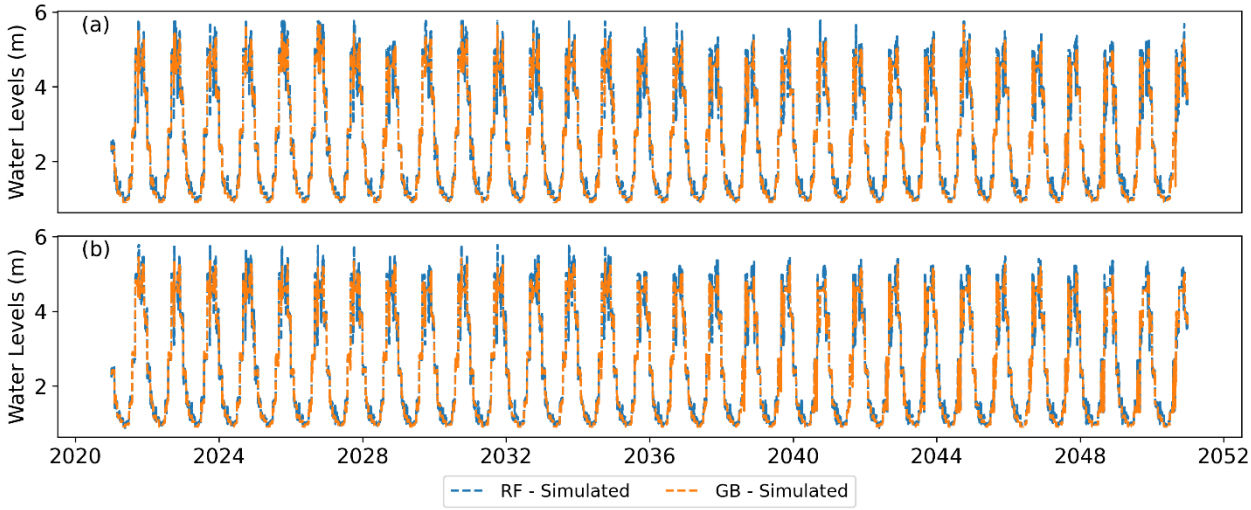

Figure 26: Simulated water levels under (a) SSP2-4.5 and (b) SSP5-8.5 scenarios using RF and GB.

6.3. Hydroclimatic event characterization

The spatial distribution maps in Figure 27 provide a clear depiction of the regional variability in hydroclimatic indices, specifically CDD, CWD, and RX1DAY, under both SSP5-8.5 (panels a, b, c) and SSP2-4.5 (panels d, e, f). These maps reveal significant spatial contrasts across the study region, emphasizing how different emissions scenarios influence hydroclimatic patterns. For instance, CWD under SSP2-4.5 showed a statistically significant breakpoint in 2041, emphasizing potential changes in wet spell patterns (Table 17, Figure 22d-f). Under SSP5-8.5, the southern parts of the study area consistently experience higher CDD values (Figure 24a), indicating prolonged dry periods. This trend aligns with the more extreme climate impacts expected in high-emissions scenarios. Conversely, under SSP2-4.5 (Figure 24d), the CDD distribution suggests a slight reduction in dry spells, with the northern regions appearing relatively less affected.

For CWD, the SSP5-8.5 scenario (Figure 27b) highlights a concentration of longer wet periods in the central regions, while the SSP2-4.5 scenario (Figure 27e) indicates a more uniform distribution of wet spells, albeit with generally shorter durations. This stark contrast between scenarios underscores the sensitivity of wet spell dynamics to emission trajectories, particularly in areas already vulnerable to hydrological variability. RX1DAY patterns (Figures 27c and 27f) also demonstrate notable spatial differences. Under SSP5-8.5, extreme daily rainfall intensities

are more pronounced in localized southern areas, suggesting a heightened risk of flash flooding and related impacts. In comparison, the SSP2-4.5 scenario shows a more balanced distribution of extreme rainfall events, with slightly reduced intensities across the region.

Table 17: Presents the p-values and breakpoints under SSP2-4.5 and SSP5-8.5.

Indices		p-value	Breakpoint
CDD	SSP2-4.5	0.059	2030
	SSP5-8.5	1	2023
CWD	SSP2-4.5	0.0337	2041
	SSP5-8.5	0.1554	2033
RX1DAY	SSP2-4.5	0.0875	2030
	SSP5-8.5	0.7184	2041

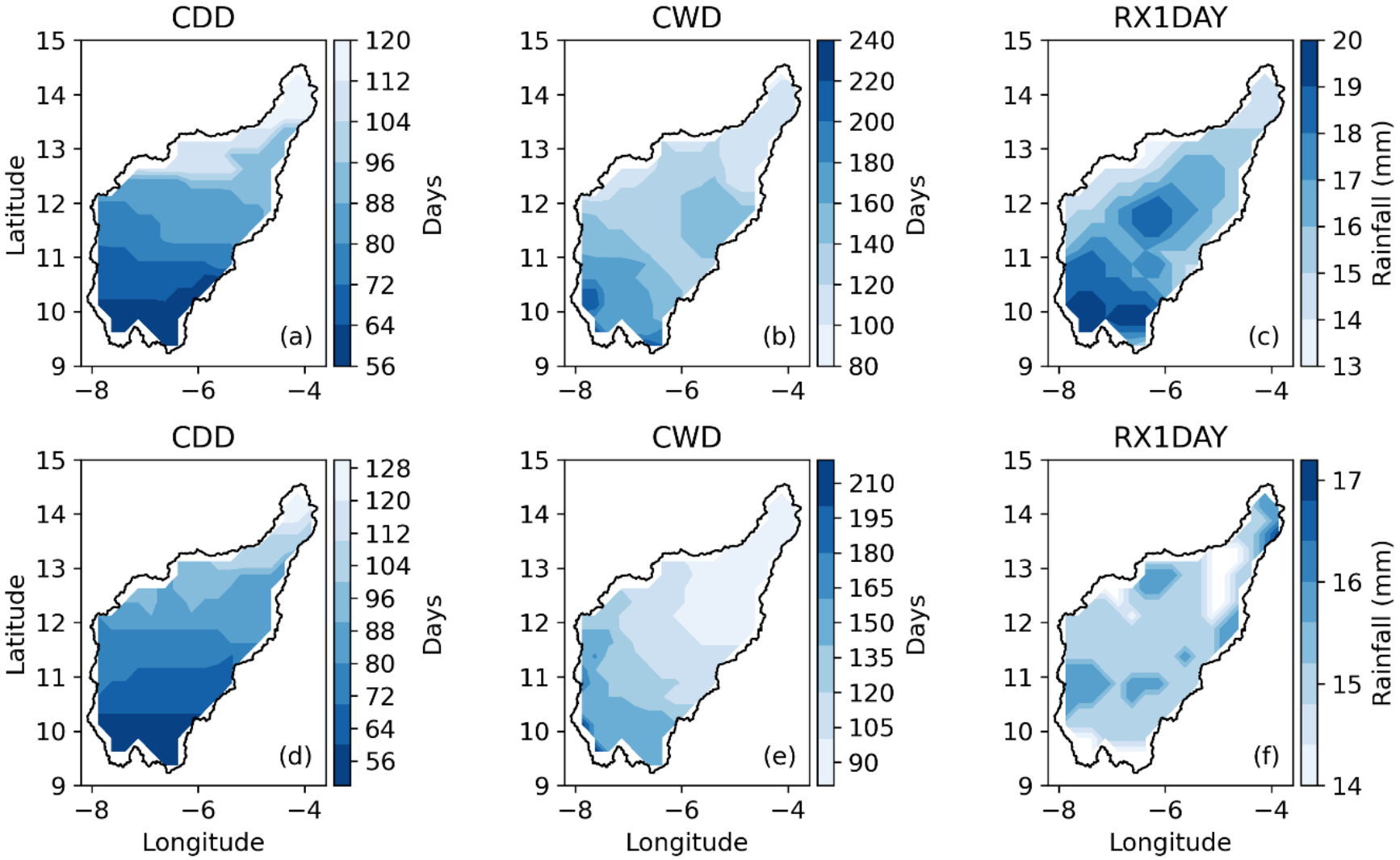

Figure 27: Spatial distribution maps for CDD, CWD, and RX1DAY. Panels (a, b, c) represent SSP5-8.5, while panels (d, e, f) correspond to SSP2-4.5.

Figure 28 provides an insightful comparison of time series trends for key hydroclimatic indices CDD, CWD, and RX1DAY under SSP2-4.5 and SSP5-8.5. Panels (a, b, c) represent the SSP2-4.5 scenario, while panels (d, e, f) correspond to SSP5-8.5. The figure highlights both similarities and notable differences in the timing of breakpoints and the direction of trends across the two scenarios.

A key observation from the analysis is that, under SSP5-8.5, the majority of indices exhibit earlier breakpoints compared to SSP2-4.5, reflecting accelerated climate impacts in the high-emissions scenario. For instance, breakpoints for CDD (Figure 28d) and CWD (Figure 28e) occur earlier in SSP5-8.5, around 2033, compared to SSP2-4.5. This pattern aligns with the expectation that more extreme greenhouse gas emissions under SSP5-8.5 will drive earlier and more rapid shifts in hydroclimatic systems. However, RX1DAY (Figures 28c and 28f) stands out as an exception to this trend. While most indices under SSP5-8.5 show earlier breakpoints, RX1DAY exhibits a delayed breakpoint in 2040, compared to its SSP2-4.5 counterpart. This divergence suggests that extreme daily rainfall intensities may take longer to display structural changes in the high-emissions scenario, highlighting the unique response of this particular index to increasing greenhouse gas concentrations.

Additionally, the trends in CWD reveal a stark contrast between the two scenarios. Under SSP2-4.5 (Figure 28b), CWD demonstrates a decreasing trend, indicating shorter wet periods over time. In contrast, under SSP5-8.5 (Figure 28e), CWD shows an increasing trend, suggesting prolonged wet conditions. This difference underscores the complex and non-linear interactions between precipitation patterns and emission scenarios, emphasizing the critical role of scenario-dependent variability in shaping future hydroclimatic regimes.

The earlier breakpoint and steeper decline in SSP5-8.5 indicate a more pronounced reduction in maximum water levels under high-emissions scenarios (Figure 29a), as compared to SSP2-4.5 (Figure 29b). The higher negative trend in SSP5-8.5, reflected by a steeper slope (-0.017 compared to -0.01 in SSP2-4.5), emphasizes the accelerated impact of higher emissions on maximum water levels (Table 18).

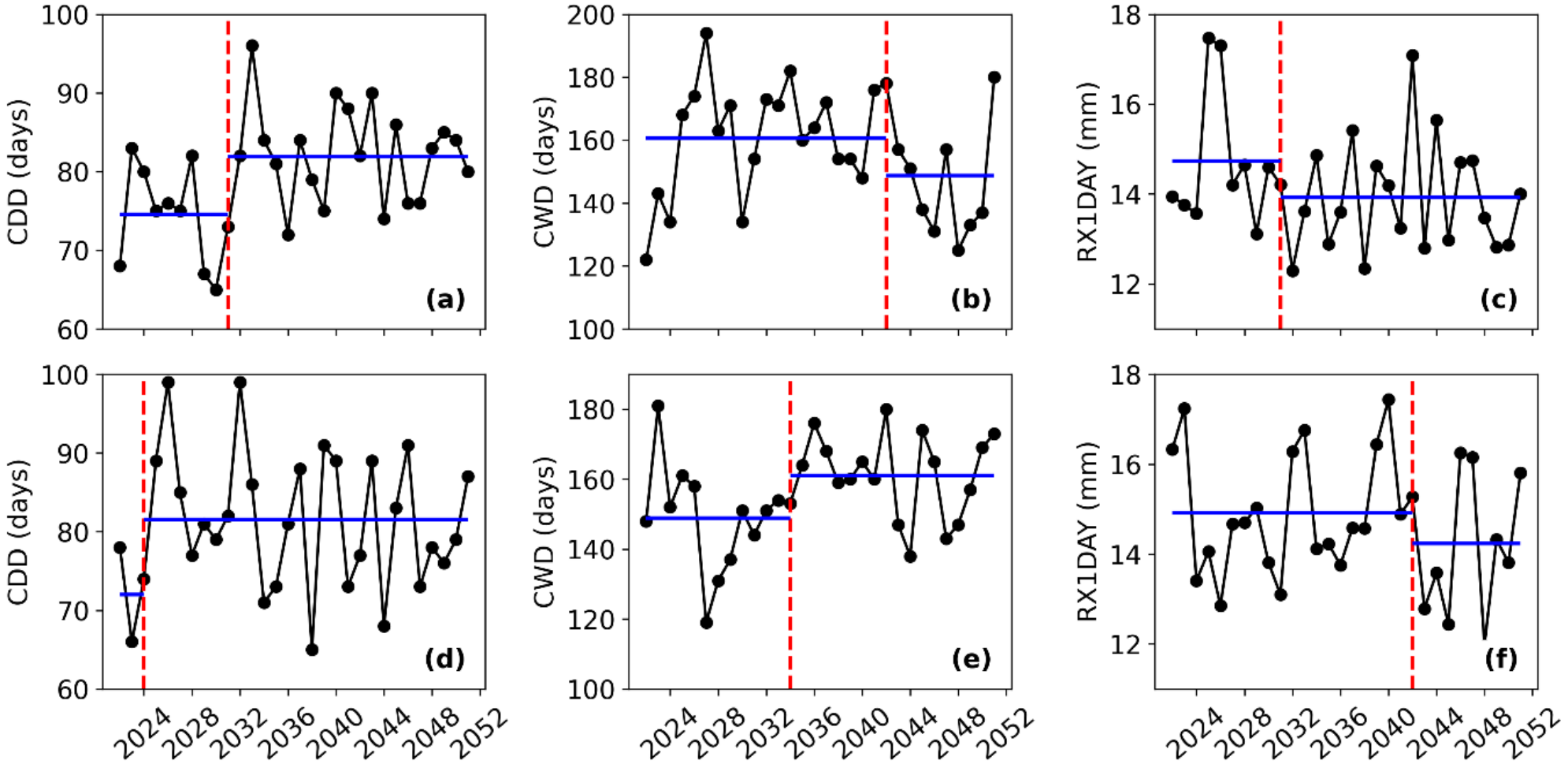

Figure 28: Time series trends for CDD, CWD, and RX1DAY. Panels (a, b, c) represent SSP2-4.5, while panels (d, e, f) correspond to SSP5-8.5.

These findings suggest a heightened vulnerability of water resources in the high-emissions pathway, with potentially severe implications for flood risk management and water availability. These findings underscore the need for robust adaptation strategies tailored to the varying impacts of different emission pathways.

Table 18: Provides insights into WLmax trends, with breakpoints identified through the Pettitt test and decreasing trends confirmed by MMK analysis.

		WLmax	
Statistical test		**SSP2-4.5**	**SSP5-8. 5**
Pettitt Test	Breakpoint	2036	2034
MMK	p-value	0.0005	0.0003
	Trend	Decreasing	
	Tau	-0.444	-0.462
	Slope	-0.01	-0.017

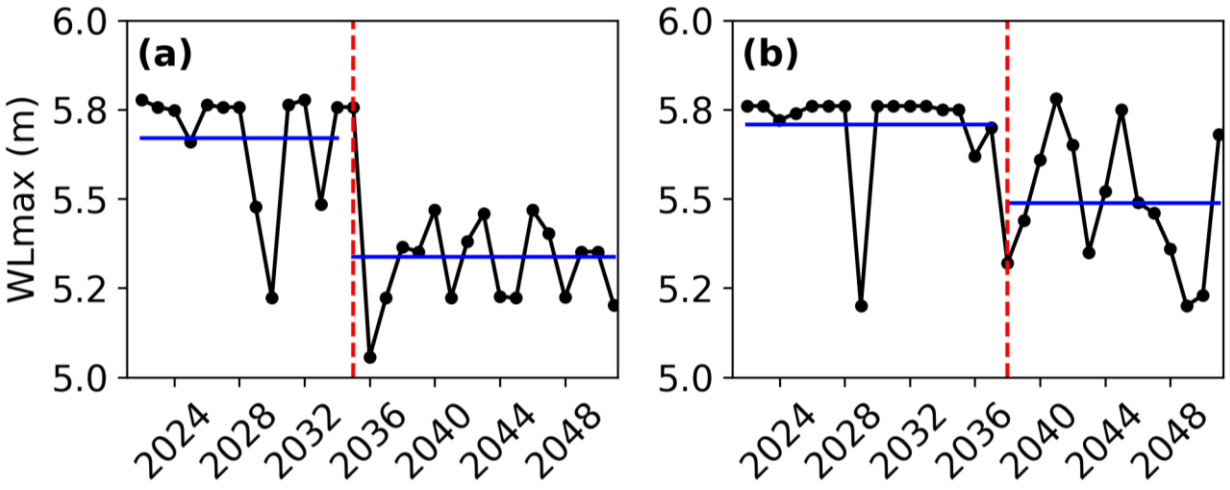

Figure 29: Future WLmax trends. Panel (a) corresponds to SSP5-8.5, while panel (b) represents SSP2-4.5.

6.4. Discussions

Our study provides a comprehensive analysis of hydroclimatic trends in the BRB, revealing significant shifts in key indices such as CDD, CWD, RX1DAY, and WLmax under different emission scenarios. The SSP2-4.5 scenario, for example, projects a significant breakpoint for CWD in 2041, indicating potential alterations in wet spell patterns critical for agriculture and water resources (Almazroui et al., 2020; Du et al., 2024). In contrast, the SSP5-8.5 scenario anticipates earlier disruptions, with a breakpoint for CWD occurring in 2033, highlighting the accelerated impacts of high-emission pathways (Sawadogo et al., 2024). While these findings

underscore the vulnerability of the BRB to climate change, certain limitations should be acknowledged. First, CMIP6 models, though state-of-the-art, can still exhibit biases when applied to regional hydrological systems, particularly in Africa where observational data are sparse (Kouyaté et al., 2025). Additionally, the scarcity of ground-based weather and hydrological monitoring in the BRB limits the ability to refine projections and fully validate the models (Muthoni et al., 2019). Addressing these data gaps is essential for enhancing predictive accuracy and ensuring the reliability of results. Nevertheless, this study makes significant contributions by integrating ML models such as RF and GB with CMIP6 outputs to provide robust future water level projections. The ability of these models to handle complex datasets and detect nonlinear trends offers valuable insights for water resource management in vulnerable regions (Xiang et al., 2021; Yaseen et al., 2019). For example, the declining WLmax trends observed under both scenarios reinforce the urgent need for adaptive strategies to mitigate flood risks and manage water scarcity (Limantol et al., 2023).

Ultimately, this work highlights the importance of combining advanced predictive tools with climate scenario analysis to address critical challenges in the BRB.

In addition to the encouraging results obtained with RF and GB, it is worth noting the growing use of recurrent neural networks, particularly Long Short-Term Memory (LSTM), for hydrological forecasting. Their ability to capture long-term temporal dependencies makes them promising tools for modeling water-level dynamics in West African basins. However, their application remains relatively novel, especially when coupled with climate projections under CMIP6 SSP scenarios (Ahmed et al., 2022b; Xiang et al., 2023). Incorporating such deep learning approaches in future research would be an interesting direction to further enhance the robustness and transferability of predictive frameworks in data-scarce and uncertainty-prone environments.

The findings serve as a foundation for informed decision-making by policymakers and stakeholders, enabling proactive measures to secure livelihoods, protect ecosystems, and enhance resilience to climate extremes (Intergovernmental Panel On Climate Change (IPCC), 2023b).

6.5. Partial conclusion

This study examines the future hydroclimatic behavior of the BRB under SSP2-4.5 and SSP5-8.5 scenarios, revealing significant changes in indices such as CWD, CDD and WLmax. For instance, wet spell patterns show notable changes around 2041 under SSP2-4.5, while the SSP5-8.5 scenario highlights earlier disruptions by 2033, emphasizing the accelerated impacts of

high-emission pathways (Sawadogo et al., 2024). The RF and GB models with CMIP6 climate projections has allowed for robust predictions, building on their proven ability to model complex, multivariate systems and handle incomplete datasets (Banadkooki et al., 2019).

Declines in WLmax, as identified by trend analyses and breakpoints in both scenarios, further highlight the vulnerability of the region's water resources. Under SSP2-4.5, significant changes in WLmax are expected by 2036, while SSP5-8.5 indicates these changes may occur as early as 2034. These findings underscore the necessity of proactive water resource management to mitigate risks, particularly in regions highly dependent on seasonal rainfall and vulnerable to climate variability (Roudier et al., 2014b).

Strategies to address these challenges include improving hydrological data collection systems, enhancing the resilience of water management infrastructure, and adopting community-driven adaptation practices. Such measures are essential for mitigating the socio-economic and ecological impacts of climate extremes, especially in regions where food security and livelihoods depend heavily on natural resources (Rameshwaran et al., 2021). By aligning predictive modeling with adaptive management strategies, stakeholders can better navigate the multifaceted challenges posed by climate change in the BRB.

Chapter 7: General conclusion and perspectives

This research set out to improve understanding of the BRB hydroclimatic variability and enhance predictive capabilities for water level fluctuations. Specifically, it addressed four major research questions involving (1) the accuracy of satellite rainfall estimates relative to ground observations, (2) the influence of changing extreme hydroclimatic events on the BRB, (3) the efficacy of machine learning models in predicting water levels, and (4) the implications of future climate scenarios for water resource availability by 2050. The findings, grounded in both statistical trend analyses and advanced machine learning (ML) approaches, offer critical insights into water resource management for this data-sparse and socio-politically challenged region.

- **Accuracy and reliability of satellite rainfall estimates**

Comparative analyses showed that satellite products particularly CHIRPS provide a promising solution for mitigating data gaps in the BRB. These findings are consistent with research in other West African basins (Houngnibo et al., 2023b; Dinku et al., 2010), where integrated satellite–ground products excel at capturing spatial and temporal rainfall variability, even in regions where conflict or logistical challenges limit the collection of ground-based data. While minor regional biases exist, CHIRPS generally outperforms other products in aligning with available gauge measurements, thereby offering a practical and reliable tool for continued monitoring in the BRB.

- **Impacts of extreme hydroclimatic events on the BRB**

Analyses of precipitation and discharge indices underscore significant spatiotemporal shifts in the BRB's hydroclimatic regime. Indicators of extreme rainfall (e.g., RX1DAY, SDII) and discharge (e.g., Qmax, Q99P) exhibit breakpoints, suggesting notable changes around the early 1990s and 2000s. These transformations parallel broader regional observations in semi-arid West Africa, where intensification of short-duration heavy rainfall and decreasing frequency of prolonged rain events have been documented (Nicholson, 2013; Sylla et al., 2015). The growing incidence of extreme flows and shifting rainfall patterns elevate flood risks in downstream wetlands, while erratic seasonal water availability threatens agricultural stability and fisheries.

- **Predictive performance of machine learning models**

Machine learning techniques such as RF and GB were effective in replicating historical water levels, handling nonlinearities and complexities better than many conventional statistical models (Anderson and Radić, 2022). Their capacity to integrate disparate data sources—

including satellite-derived precipitation and incomplete in-situ observations—makes them well-suited for the BRB, where robust ground-based data are often lacking. Among the models tested, both RF and GB consistently performed well, with slight performance advantages sometimes observed for SVM under highly variable datasets (Saxena et al., 2023). However, tuning RF and GB further can achieve comparably stable performance. These results validate the potential of ML-based tools to enhance real-time forecasting and long-term planning in data-sparse basins.

- **Future water levels under SSP2-4.5 and SSP5-8.5 scenarios**

Projections incorporating CMIP6 outputs highlight distinct pathways for water availability in the BRB by 2050. Under the more moderate SSP2-4.5 scenario, changes in wet and dry spell patterns manifest more gradually, with breakpoints appearing around the early to mid-2040s (Sawadogo et al., 2024). By contrast, SSP5-8.5 a high-emission scenario signals earlier and more pronounced disruptions, including potential declines in maximum water levels (WLmax) by the mid-2030s. These findings align with broader regional studies forecasting substantial reductions in water availability and more frequent hydroclimatic extremes across West Africa (Sylla et al., 2018; Rameshwaran et al., 2021). They underscore the urgent need for adaptive water management, infrastructure resilience, and policy interventions to mitigate the compounded effects of climate change, social unrest, and infrastructural limitations.

Implications and Perspectives

Collectively, these results confirm that satellite rainfall products can successfully bridge observational gaps, offering a feasible data source where traditional monitoring faces socio-political challenge. The observed shifts in extreme events underscore the necessity of proactive planning, both to manage flood risks and secure agricultural water supplies. ML models, proven effective in predicting water-level fluctuations, should form the backbone of emerging early warning systems and long-term forecasting strategies. Finally, the divergent futures outlined by SSP2-4.5 and SSP5-8.5 highlight the importance of integrating climate scenario planning into all levels of basin management.

Building on the findings of this research, future investigations could prioritize the downscaling of global climate models to achieve higher spatial resolution, thereby capturing local microclimatic variations in the BRB more accurately. Addressing the challenge of limited ground-based observation networks, particularly in regions affected by insecurity, remains critical. Enhancing these networks would significantly improve the calibration and validation

processes for both satellite rainfall estimates and machine learning (ML) models, boosting their predictive accuracy.

Moreover, integrating socio-economic dimensions into future studies would enable a more holistic approach to water resource management. Such integration could provide critical insights into water allocation dynamics, ensuring equitable distribution across different societal groups and informing policy reforms that consider both environmental and social needs. The insights derived from this study lay a solid foundation for the development of adaptive, data-driven, and sustainable water management strategies in the BRB. These findings are particularly relevant for stakeholders aiming to implement resilient water governance frameworks capable of addressing the compounded challenges posed by climate variability, socio-political instability, and infrastructural constraints.

8. References

Addor, N., Do, H.X., Alvarez-Garreton, C., Coxon, G., Fowler, K., Mendoza, P.A., 2020. Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges. Hydrological Sciences Journal 65, 712–725. https://doi.org/10.1080/02626667.2019.1683182

Adounkpe, J.Y.P., Diallo, B.A.A., Alamou, E.A., Ali, A., n.d. Predicting Discharge in Catchment Outlet Using Deep Learning: Case Study of the Ansongo-Niamey Basin 7.

Ahmed, A.N., Yafouz, A., Birima, A.H., Kisi, O., Huang, Y.F., Sherif, M., Sefelnasr, A., El-Shafie, A., 2022a. Water level prediction using various machine learning algorithms: a case study of Durian Tunggal river, Malaysia. Engineering Applications of Computational Fluid Mechanics 16, 422–440. https://doi.org/10.1080/19942060.2021.2019128

Ahmed, A.N., Yafouz, A., Birima, A.H., Kisi, O., Huang, Y.F., Sherif, M., Sefelnasr, A., El-Shafie, A., 2022b. Water level prediction using various machine learning algorithms: a case study of Durian Tunggal river, Malaysia. Engineering Applications of Computational Fluid Mechanics 16, 422–440. https://doi.org/10.1080/19942060.2021.2019128

Allawi, M.F., Jaafar, O., Mohamad Hamzah, F., Abdullah, S.M.S., El-shafie, A., 2018. Review on applications of artificial intelligence methods for dam and reservoir-hydro-environment models. Environ Sci Pollut Res 25, 13446–13469. https://doi.org/10.1007/s11356-018-1867-8

Almazroui, M., Saeed, F., Saeed, S., Nazrul Islam, M., Ismail, M., Klutse, N.A.B., Siddiqui, M.H., 2020. Projected Change in Temperature and Precipitation Over Africa from CMIP6. Earth Syst Environ 4, 455–475. https://doi.org/10.1007/s41748-020-00161-x

Amogu, O., Descroix, L., Yéro, K.S., Le Breton, E., Mamadou, I., Ali, A., Vischel, T., Bader, J.-C., Moussa, I.B., Gautier, E., Boubkraoui, S., Belleudy, P., 2010. Increasing River Flows in the Sahel? Water 2, 170–199. https://doi.org/10.3390/w2020170

Anderson, S., Radić, V., 2022. Evaluation and interpretation of convolutional long short-term memory networks for regional hydrological modelling. Hydrol. Earth Syst. Sci. 26, 795–825. https://doi.org/10.5194/hess-26-795-2022

Andersson, J.C.M., Arheimer, B., Traoré, F., Gustafsson, D., Ali, A., 2017. Process refinements improve a hydrological model concept applied to the Niger River basin. Hydrological Processes 31, 4540–4554. https://doi.org/10.1002/hyp.11376

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B.R., Prat, O.P., 2015. PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. Bulletin of the American Meteorological Society 96, 69–83. https://doi.org/10.1175/BAMS-D-13-00068.1

Ba, B., 2023. Conflicts between Farmers and Livestock Breeders in Mali: A Review of Dynamics and Issues in the Mopti Region.

Balcázar, L., Bâ, K.M., Díaz-Delgado, C., Gómez-Albores, M.A., Gaona, G., Minga-León, S., 2022. Development and Assessment of Seasonal Rainfall Forecasting Models for the Bani and the Senegal Basins by Identifying the Best Predictive Teleconnection. Remote Sensing 14, 6397. https://doi.org/10.3390/rs14246397

Bamal, A., Uddin, M.G., Olbert, A.I., 2024. Harnessing machine learning for assessing climate change influences on groundwater resources: A comprehensive review. Heliyon 10, e37073. https://doi.org/10.1016/j.heliyon.2024.e37073

Banadkooki, F.B., Ehteram, M., Ahmed, A.N., Fai, C.M., Afan, H.A., Ridwam, W.M., Sefelnasr, A., El-Shafie, A., 2019. Precipitation Forecasting Using Multilayer Neural Network and Support Vector Machine Optimization Based on Flow Regime Algorithm Taking into Account Uncertainties of Soft Computing Models. Sustainability 11, 6681. https://doi.org/10.3390/su11236681

Basse, J., Sabaly, H.N., Diba, I., Sarr, A.B., Camara, M., Diedhiou, A., 2021. Synoptic variability associated with wet and dry spells in the Western Sahel. Theor Appl Climatol 146, 1019–1029. https://doi.org/10.1007/s00704-021-03774-y

Biasutti, M., 2019. Rainfall trends in the African Sahel: Characteristics, processes, and causes. WIREs Climate Change 10, e591. https://doi.org/10.1002/wcc.591

Bodian, A., Dezetter, A., Diop, L., Deme, A., Djaman, K., Diop, A., 2018. Future Climate Change Impacts on Streamflows of Two Main West Africa River Basins: Senegal and Gambia. Hydrology 5, 21. https://doi.org/10.3390/hydrology5010021

Bodian, A., Diop, L., Panthou, G., Dacosta, H., Deme, A., Dezetter, A., Ndiaye, P.M., Diouf, I., Vischel, T., 2020. Recent Trend in Hydroclimatic Conditions in the Senegal River Basin. Water 12, 436. https://doi.org/10.3390/w12020436

Bostan, P.A., Heuvelink, G.B.M., Akyurek, S.Z., 2012. Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey. International Journal of Applied Earth Observation and Geoinformation 19, 115–126. https://doi.org/10.1016/j.jag.2012.04.010

Burn, D.H., Hag Elnur, M.A., 2002. Detection of hydrologic trends and variability. Journal of Hydrology 255, 107–122. https://doi.org/10.1016/S0022-1694(01)00514-5

Carr, T.W., Mkuhlani, S., Segnon, A.C., Ali, Z., Zougmoré, R., Dangour, A.D., Green, R., Scheelbeek, P., 2022. Climate change impacts and adaptation strategies for crops in West Africa: a systematic review. Environ. Res. Lett. 17, 053001. https://doi.org/10.1088/1748-9326/ac61c8

Chaibou Begou, J., Jomaa, S., Benabdallah, S., Bazie, P., Afouda, A., Rode, M., 2016. Multi-Site Validation of the SWAT Model on the Bani Catchment: Model Performance and Predictive Uncertainty. Water 8, 178. https://doi.org/10.3390/w8050178

Coron, L., Perrin, C., Delaigue, O., Andréassian, V., Thirel, G., n.d. airGR: a suite of lumped hydrological models in an R-package 1.

De Luca, P., Messori, G., Wilby, R.L., Mazzoleni, M., Di Baldassarre, G., 2020. Concurrent wet and dry hydrological extremes at the global scale. Earth Syst. Dynam. 11, 251–266. https://doi.org/10.5194/esd-11-251-2020

Dembélé, M., Zwart, S.J., 2016. Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa. International Journal of Remote Sensing 37, 3995–4014. https://doi.org/10.1080/01431161.2016.1207258

Demirel, M.C., Venancio, A., Kahya, E., 2009. Flow forecast by SWAT model and ANN in Pracana basin, Portugal. Advances in Engineering Software 40, 467–473. https://doi.org/10.1016/j.advengsoft.2008.08.002

Descroix, L., Genthon, P., Amogu, O., Rajot, J.-L., Sighomnou, D., Vauclin, M., 2012. Change in Sahelian Rivers hydrograph: The case of recent red floods of the Niger River in the Niamey region. Global and Planetary Change 98–99, 18–30. https://doi.org/10.1016/j.gloplacha.2012.07.009

Di Baldassarre, G., Sivapalan, M., Rusca, M., Cudennec, C., Garcia, M., Kreibich, H., Konar, M., Mondino, E., Mård, J., Pande, S., Sanderson, M.R., Tian, F., Viglione, A., Wei, J., Wei, Y., Yu, D.J., Srinivasan, V., Blöschl, G., 2019. Sociohydrology: Scientific Challenges in Addressing the Sustainable Development Goals. Water Resources Research 55, 6327–6355. https://doi.org/10.1029/2018WR023901

Dinku, T., Ceccato, P., Cressman, K., Connor, S.J., 2010. Evaluating Detection Skills of Satellite Rainfall Estimates over Desert Locust Recession Regions. Journal of Applied Meteorology and Climatology 49, 1322–1332. https://doi.org/10.1175/2010JAMC2281.1

Diop, L., Samadianfard, S., Bodian, A., Yaseen, Z.M., Ghorbani, M.A., Salimi, H., 2020. Annual Rainfall Forecasting Using Hybrid Artificial Intelligence Model: Integration of Multilayer Perceptron with Whale Optimization Algorithm. Water Resour Manage 34, 733–746. https://doi.org/10.1007/s11269-019-02473-8

Dos Santos, M.R.W., 2023. Water cooperation within West Africa's major transboundary river basins. Regions and Cohesion 13, 25–52. https://doi.org/10.3167/reco.2023.130203

Du, H., Tan, M.L., Zhang, F., Chun, K.P., Li, L., Kabir, M.H., 2024. Evaluating the effectiveness of CHIRPS data for hydroclimatic studies. Theor Appl Climatol 155, 1519–1539. https://doi.org/10.1007/s00704-023-04721-9

Ehsanzadeh, E., Adamowski, K., 2010. Trends in timing of low stream flows in Canada: impact of autocorrelation and long-term persistence. Hydrological Processes 24, 970–980. https://doi.org/10.1002/hyp.7533

Fofana, M., Adounkpe, J., Larbi, I., Hounkpe, J., Djan'na Koubodana, H., Toure, A., Bokar, H., Dotse, S.-Q., Limantol, A.M., 2022. Urban flash flood and extreme rainfall events trend analysis in Bamako, Mali. Environmental Challenges 6, 100449. https://doi.org/10.1016/j.envc.2022.100449

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci Data 2, 150066. https://doi.org/10.1038/sdata.2015.66

Ghazvinian, H., Mousavi, S.-F., Karami, H., Farzin, S., Ehteram, M., Hossain, M.S., Fai, C.M., Hashim, H.B., Singh, V.P., Ros, F.C., Ahmed, A.N., Afan, H.A., Lai, S.H., El-Shafie, A., 2019. Integrated support vector regression and an improved particle swarm optimization-based model for solar radiation prediction. PLOS ONE 14, e0217634. https://doi.org/10.1371/journal.pone.0217634

Guan, X., Zhang, J., Yang, Q., Tang, X., Liu, C., Jin, J., Liu, Y., Bao, Z., Wang, G., 2020. Evaluation of Precipitation Products by Using Multiple Hydrological Models over the Upper Yellow River Basin, China. Remote Sensing 12, 4023. https://doi.org/10.3390/rs12244023

Hamed, K.H., Ramachandra Rao, A., 1998. A modified Mann-Kendall trend test for autocorrelated data. Journal of Hydrology 204, 182–196. https://doi.org/10.1016/S0022-1694(97)00125-X

Hamitouche, Y., Zeroual, A., Meddi, M., Assani, A.A., Alkama, R., Şen, Z., Zhang, X., 2024. Projected Changes in Extreme Precipitation Patterns across Algerian Sub-Regions. Water 16, 1353. https://doi.org/10.3390/w16101353

Hassani, A., Azapagic, A., Shokri, N., 2020. Predicting long-term dynamics of soil salinity and sodicity on a global scale. Proc. Natl. Acad. Sci. U.S.A. 117, 33017–33027. https://doi.org/10.1073/pnas.2013771117

Hassani, A., Smith, P., Shokri, N., 2024. Negative correlation between soil salinity and soil organic carbon variability. Proc. Natl. Acad. Sci. U.S.A. 121, e2317332121. https://doi.org/10.1073/pnas.2317332121

Hipni, A., El-shafie, A., Najah, A., Karim, O.A., Hussain, A., Mukhlisin, M., 2013. Daily Forecasting of Dam Water Levels: Comparing a Support Vector Machine (SVM) Model With Adaptive Neuro Fuzzy Inference System (ANFIS). Water Resour Manage 27, 3803–3823. https://doi.org/10.1007/s11269-013-0382-4

Houngnibo, M.C.M., Minoungou, B., Traore, S.B., Maidment, R.I., Alhassane, A., Ali, A., 2023a. Validation of high-resolution satellite precipitation products over West Africa for rainfall monitoring and early warning. Front. Clim. 5. https://doi.org/10.3389/fclim.2023.1185754

Houngnibo, M.C.M., Minoungou, B., Traore, S.B., Maidment, R.I., Alhassane, A., Ali, A., 2023b. Validation of high-resolution satellite precipitation products over West Africa for rainfall monitoring and early warning. Front. Clim. 5, 1185754. https://doi.org/10.3389/fclim.2023.1185754

Houngnibo, M.C.M., Minoungou, B., Traore, S.B., Maidment, R.I., Alhassane, A., Ali, A., 2023c. Validation of high-resolution satellite precipitation products over West Africa for rainfall monitoring and early warning. Front. Clim. 5. https://doi.org/10.3389/fclim.2023.1185754

Ibrahim, M., Wisser, D., Ali, A., Diekkrüger, B., Seidou, O., Mariko, A., Afouda, A., 2017. Water Balance Analysis over the Niger Inland Delta-Mali: Spatio-Temporal Dynamics of the Flooded Area and Water Losses. Hydrology 4, 40. https://doi.org/10.3390/hydrology4030040

Intergovernmental Panel On Climate Change (IPCC), 2023a. Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 1st ed. Cambridge University Press. https://doi.org/10.1017/9781009157896

Intergovernmental Panel On Climate Change (IPCC), 2023b. Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 1st ed. Cambridge University Press. https://doi.org/10.1017/9781009325844

Jeong, H.-G., Ahn, J.-B., Lee, J., Shim, K.-M., Jung, M.-P., 2020. Improvement of daily precipitation estimations using PRISM with inverse-distance weighting. Theor Appl Climatol 139, 923–934. https://doi.org/10.1007/s00704-019-03012-6

Jing, L., Qingwen, M., Wenhua, L., Yanying, B., C, D.B.G., Zheng, Y., 2014. Spatial Variability Analysis of Soil Nutrients Based on GIS and Geostatistics: A Case Study of Yisa Township, Yunnan, China. Journal of Resources and Ecology 5, 348–355. https://doi.org/10.5814/j.issn.1674-764x.2014.04.010

Kamel, A.H., Afan, H.A., Sherif, M., Ahmed, A.N., El-Shafie, A., 2021. RBFNN versus GRNN modeling approach for sub-surface evaporation rate prediction in arid region. Sustainable Computing: Informatics and Systems 30, 100514. https://doi.org/10.1016/j.suscom.2021.100514

Kasei, R., Diekkrüger, B., Leemhuis, C., 2010. Drought frequency in the Volta Basin of West Africa. Sustain Sci 5, 89–97. https://doi.org/10.1007/s11625-009-0101-5

Katipoğlu, O.M., 2022. Spatial analysis of seasonal precipitation using various interpolation methods in the Euphrates basin, Turkey. Acta Geophys. 70, 859–878. https://doi.org/10.1007/s11600-022-00756-0

Khaliq, M.N., Ouarda, T.B.M.J., Gachon, P., Sushama, L., St-Hilaire, A., 2009. Identification of hydrological trends in the presence of serial and cross correlations: A review of selected methods and their application to annual flow regimes of Canadian rivers. Journal of Hydrology 368, 117–130. https://doi.org/10.1016/j.jhydrol.2009.01.035

Kohnert, D., 2023. Is Water a Blessing or a Curse? How to Address Water Conflicts in West Africa. https://doi.org/10.2139/ssrn.4526004

Kouyaté, F., Guédjé, F.K., Ndiaye, A., Ganni Mampo, O.M., 2025. Spatial and Temporal Variability of Extreme Hydroclimatic Events in the Bani River Basin. Hydrology 12, 5. https://doi.org/10.3390/hydrology12010005

Kumar, P., Debele, S.E., Sahani, J., Rawat, N., Marti-Cardona, B., Alfieri, S.M., Basu, B., Basu, A.S., Bowyer, P., Charizopoulos, N., Gallotti, G., Jaakko, J., Leo, L.S., Loupis, M., Menenti, M., Mickovski, S.B., Mun, S.-J., Gonzalez-Ollauri, A., Pfeiffer, J., Pilla, F., Pröll, J., Rutzinger, M., Santo, M.A., Sannigrahi, S., Spyrou, C., Tuomenvirta, H., Zieher, T., 2021. Nature-based solutions efficiency evaluation against natural hazards:

Modelling methods, advantages and limitations. Science of The Total Environment 784, 147058. https://doi.org/10.1016/j.scitotenv.2021.147058

Kundzewicz, Z.W., Robson, A.J., 2004. Change detection in hydrological records—a review of the methodology / Revue méthodologique de la détection de changements dans les chroniques hydrologiques. Hydrological Sciences Journal 49, 7–19. https://doi.org/10.1623/hysj.49.1.7.53993

Limantol, A.M., Larbi, I., Dotse, S.-Q., Okafor, G.C., Asare-Nuamah, P., Frimpong, L.K., Alhassan, A.-R.M., Angmor, E., Joel Premkumar, J.P., Lutterodt, G., Asiamah, T.A., Atta-Darkwa, T., Kumi, N., Prempeh, N.A., 2023. An increase in temperature under the shared socioeconomic scenarios in the Volta River Basin, West Africa: implications for economic development. Journal of Water and Climate Change 14, 2808–2824. https://doi.org/10.2166/wcc.2023.141

Linard, C., Gilbert, M., Snow, R.W., Noor, A.M., Tatem, A.J., 2012. Population Distribution, Settlement Patterns and Accessibility across Africa in 2010. PLoS ONE 7, e31743. https://doi.org/10.1371/journal.pone.0031743

Louvet, S., Paturel, J.E., Mahé, G., Rouché, N., Koité, M., 2016. Comparison of the spatiotemporal variability of rainfall from four different interpolation methods and impact on the result of GR2M hydrological modeling—case of Bani River in Mali, West Africa. Theor Appl Climatol 123, 303–319. https://doi.org/10.1007/s00704-014-1357-y

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing 21, 1303–1330. https://doi.org/10.1080/014311600210191

Mahe, G., 2009. Surface/groundwater interactions in the Bani and Nakambe rivers, tributaries of the Niger and Volta basins, West Africa. Hydrological Sciences Journal 54, 704–712. https://doi.org/10.1623/hysj.54.4.704

Mahe, G., Lienou, G., Descroix, L., Bamba, F., Paturel, J.E., Laraque, A., Meddi, M., Habaieb, H., Adeaga, O., Dieulin, C., Chahnez Kotti, F., Khomsi, K., 2013. The rivers of Africa: witness of climate change and human impact on the environment. Hydrological Processes 27, 2105–2114. https://doi.org/10.1002/hyp.9813

Maidment, R.I., Grimes, D., Black, E., Tarnavsky, E., Young, M., Greatrex, H., Allan, R.P., Stein, T., Nkonde, E., Senkunda, S., Alcántara, E.M.U., 2017. A new, long-term daily

satellite-based rainfall dataset for operational monitoring in Africa. Sci Data 4, 170063. https://doi.org/10.1038/sdata.2017.63

Mishra, A.K., Singh, V.P., 2011. Drought modeling – A review. Journal of Hydrology 403, 157–175. https://doi.org/10.1016/j.jhydrol.2011.03.049

Muthoni, F.K., Odongo, V.O., Ochieng, J., Mugalavai, E.M., Mourice, S.K., Hoesche-Zeledon, I., Mwila, M., Bekunda, M., 2019. Long-term spatial-temporal trends and variability of rainfall over Eastern and Southern Africa. Theor Appl Climatol 137, 1869–1882. https://doi.org/10.1007/s00704-018-2712-1

Ndiaye, A., Mbaye, M.L., Arnault, J., Camara, M., Lawin, A.E., 2023. Characterization of Extreme Rainfall and River Discharge over the Senegal River Basin from 1982 to 2021. Hydrology 10, 204. https://doi.org/10.3390/hydrology10100204

Nicholson, S.E., 2013. The West African Sahel: A Review of Recent Studies on the Rainfall Regime and Its Interannual Variability. ISRN Meteorology 2013, 1–32. https://doi.org/10.1155/2013/453521

O'Brien, G.C., Dickens, C., Baker, C., Stassen, R., van Weert, F., 2020. Sustainable Floodplains: Linking E-Flows to Floodplain Management, Ecosystems, and Livelihoods in the Sahel of North Africa. Sustainability 12, 10578. https://doi.org/10.3390/su122410578

Ogunrinde, A.T., Oguntunde, P.G., Akinwumiju, A.S., Fasinmirin, J.T., 2021. Evaluation of the impact of climate change on the characteristics of drought in Sahel Region of Nigeria: 1971–2060. African Geographical Review 40, 192–210. https://doi.org/10.1080/19376812.2020.1814826

Okoronkwo, D.J., Ozioko, R.I., Ugwoke, R.U., Nwagbo, U.V., Nwobodo, C., Ugwu, C.H., Okoro, G.G., Mbah, E.C., 2024. Climate smart agriculture? Adaptation strategies of traditional agriculture to climate change in sub-Saharan Africa. Front. Clim. 6, 1272320. https://doi.org/10.3389/fclim.2024.1272320

Oliver, M.A., Webster, R., 2014. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. CATENA 113, 56–69. https://doi.org/10.1016/j.catena.2013.09.006

Onuoha, F.C., Ojewale, O., Akogwu, C.J., 2023. Climate Change and Natural Resource Conflict in ECOW AS and ECCAS Regions: Implications for State Security Forces. African Journal on Conflict Resolution 23, 7–36. https://doi.org/10.10520/ejc-accordr_v23_n2_a2

Osman, A.I., Hefny, M., Abdel Maksoud, M.I.A., Elgarahy, A.M., Rooney, D.W., 2021. Recent advances in carbon capture storage and utilisation technologies: a review. Environ Chem Lett 19, 797–849. https://doi.org/10.1007/s10311-020-01133-3

Oyerinde, G., Hountondji, F., Lawin, A., Odofin, A., Afouda, A., Diekkrüger, B., 2017. Improving Hydro-Climatic Projections with Bias-Correction in Sahelian Niger Basin, West Africa. Climate 5, 8. https://doi.org/10.3390/cli5010008

Oyerinde, G.T., Fademi, I.O., Denton, O.A., 2017. Modeling runoff with satellite-based rainfall estimates in the Niger basin. Cogent Food & Agriculture 3, 1363340. https://doi.org/10.1080/23311932.2017.1363340

Paturel, J.-E., 2014. Exercice de scénarisation hydrologique en Afrique de l'Ouest—Bassin du Bani. Hydrological Sciences Journal 59, 1135–1153. https://doi.org/10.1080/02626667.2013.834340

Paturel', J.-E., DlAWARA, A., Siou, L.K.A., Talin, E., FERRy, L., Mahe', G., n.d. Caractérisation de la sécheresse hydropluviométrique du Bani, principal affluent du fleuve Niger au Mali 10.

Peters, K., Dupar, M., Opitz-Stapleton, S., Lovell, E., Budimir, M., Brown, S., Cao, Y., 2020. Climate change, conflict and fragility: an evidence review and recommendations for research and action (Report). Overseas Development Institute.

Pettitt, A.N., 1979. A Non-Parametric Approach to the Change-Point Problem. Applied Statistics 28, 126. https://doi.org/10.2307/2346729

Pham, L.T., Luo, L., Finley, A., 2021. Evaluation of random forests for short-term daily streamflow forecasting in rainfall- and snowmelt-driven watersheds. Hydrol. Earth Syst. Sci. 25, 2997–3015. https://doi.org/10.5194/hess-25-2997-2021

Picouet, C., Hingray, B., Olivry, J., 2000. Modélisation conceptuelle globale du régime du transport particulaire sur les fleuves tropicaux d'Afrique : application aux bassins du Niger supérieur et du Bani, Mali. rseau 13, 463–481. https://doi.org/10.7202/705403ar

Quenum, G.M.L.D., Nkrumah, F., Klutse, N.A.B., Sylla, M.B., 2021. Spatiotemporal Changes in Temperature and Precipitation in West Africa. Part I: Analysis with the CMIP6 Historical Dataset. Water 13, 3506. https://doi.org/10.3390/w13243506

Rameshwaran, P., Bell, V.A., Davies, H.N., Kay, A.L., 2021. How might climate change affect river flows across West Africa? Climatic Change 169, 21. https://doi.org/10.1007/s10584-021-03256-0

Rezaei, S., Ranjineh Khojasteh, E., Faridazad, M., 2020. Improving geostatistical predictions of two environmental variables using Bayesian maximum entropy in the Sungun mining

site. Stoch Environ Res Risk Assess 34, 1775–1794. https://doi.org/10.1007/s00477-020-01863-4

Ridwan, W.M., Sapitang, M., Aziz, A., Kushiar, K.F., Ahmed, A.N., El-Shafie, A., 2021. Rainfall forecasting model using machine learning methods: Case study Terengganu, Malaysia. Ain Shams Engineering Journal 12, 1651–1663. https://doi.org/10.1016/j.asej.2020.09.011

Roudier, P., Ducharne, A., Feyen, L., 2014a. Climate change impacts on runoff in West Africa: a review. Hydrol. Earth Syst. Sci. 18, 2789–2801. https://doi.org/10.5194/hess-18-2789-2014

Roudier, P., Ducharne, A., Feyen, L., 2014b. Climate change impacts on runoff in West Africa: a review. Hydrol. Earth Syst. Sci. 18, 2789–2801. https://doi.org/10.5194/hess-18-2789-2014

Roudier, P., Mahe, G., 2010a. Study of water stress and droughts with indicators using daily data on the Bani river (Niger basin, Mali). Intl Journal of Climatology 30, 1689–1705. https://doi.org/10.1002/joc.2013

Roudier, P., Mahe, G., 2010b. Study of water stress and droughts with indicators using daily data on the Bani river (Niger basin, Mali). Intl Journal of Climatology 30, 1689–1705. https://doi.org/10.1002/joc.2013

Ruelland, D., Ardoin-Bardin, S., Collet, L., Roucou, P., 2012. Simulating future trends in hydrological regime of a large Sudano-Sahelian catchment under climate change. Journal of Hydrology 424–425, 207–216. https://doi.org/10.1016/j.jhydrol.2012.01.002

Sadeghi, M., Asanjan, A.A., Faridzad, M., Nguyen, P., Hsu, K., Sorooshian, S., Braithwaite, D., 2019. PERSIANN-CNN: Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Convolutional Neural Networks. Journal of Hydrometeorology 20, 2273–2289. https://doi.org/10.1175/JHM-D-19-0110.1

Saldanha Do Carmo, C., Silventoinen, P., Nordgård, C.T., Poudroux, C., Dessev, T., Zobel, H., Holtekjølen, A.K., Draget, K.I., Holopainen-Mantila, U., Knutsen, S.H., Sahlstrøm, S., 2020. Is dehulling of peas and faba beans necessary prior to dry fractionation for the production of protein- and starch-rich fractions? Impact on physical properties, chemical composition and techno-functional properties. Journal of Food Engineering 278, 109937. https://doi.org/10.1016/j.jfoodeng.2020.109937

Sang, Y.-F., Wang, Z., Liu, C., 2014. Comparison of the MK test and EMD method for trend identification in hydrological time series. Journal of Hydrology 510, 293–298. https://doi.org/10.1016/j.jhydrol.2013.12.039

Sapitang, M., M. Ridwan, W., Faizal Kushiar, K., Najah Ahmed, A., El-Shafie, A., 2020. Machine Learning Application in Reservoir Water Level Forecasting for Sustainable Hydropower Generation Strategy. Sustainability 12, 6121. https://doi.org/10.3390/su12156121

Sawadogo, W., Neya, T., Semde, I., Korahiré, J.A., Combasséré, A., Traoré, D.E., Ouedraogo, P., Diasso, U.J., Abiodun, B.J., Bliefernicht, J., Kunstmann, H., 2024. Potential impacts of climate change on the sudan-sahel region in West Africa – Insights from Burkina Faso. Environmental Challenges 15, 100860. https://doi.org/10.1016/j.envc.2024.100860

Saxena, S., Herrmann, C., Hur, J., Kar, A., Norouzi, M., Sun, D., Fleet, D.J., 2023. The Surprising Effectiveness of Diffusion Models for Optical Flow and Monocular Depth Estimation. Advances in Neural Information Processing Systems 36, 39443–39469.

Serinaldi, F., Kilsby, C.G., 2016. The importance of prewhitening in change point analysis under persistence. Stoch Environ Res Risk Assess 30, 763–777. https://doi.org/10.1007/s00477-015-1041-5

Shahabi, H., Shirzadi, A., Ghaderi, K., Omidvar, E., Al-Ansari, N., Clague, J.J., Geertsema, M., Khosravi, K., Amini, A., Bahrami, S., Rahmati, O., Habibi, K., Mohammadi, A., Nguyen, H., Melesse, A.M., Ahmad, B.B., Ahmad, A., 2020. Flood Detection and Susceptibility Mapping Using Sentinel-1 Remote Sensing Data and a Machine Learning Approach: Hybrid Intelligence of Bagging Ensemble Based on K-Nearest Neighbor Classifier. Remote Sensing 12, 266. https://doi.org/10.3390/rs12020266

Shahriari, B., Bouchard-Cote, A., de Freitas, N., n.d. Unbounded Bayesian Optimization via Regularization.

Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical Bayesian Optimization of Machine Learning Algorithms. https://doi.org/10.48550/arXiv.1206.2944

Sultan, B., Gaetani, M., 2016. Agriculture in West Africa in the Twenty-First Century: Climate Change and Impacts Scenarios, and Potential for Adaptation. Front. Plant Sci. 7. https://doi.org/10.3389/fpls.2016.01262

Sylla, M.B., Faye, A., Klutse, N.A.B., Dimobe, K., 2018. Projected increased risk of water deficit over major West African river basins under future climates. Climatic Change 151, 247–258. https://doi.org/10.1007/s10584-018-2308-x

Sylla, M.B., Giorgi, F., Pal, J.S., Gibba, P., Kebe, I., Nikiema, M., 2015. Projected Changes in the Annual Cycle of High-Intensity Precipitation Events over West Africa for the Late

Twenty-First Century*. Journal of Climate 28, 6475–6488. https://doi.org/10.1175/JCLI-D-14-00854.1

Tahmasebi, P., Sahimi, M., 2021. Special issue on machine learning for water resources and subsurface systems. Advances in Water Resources 149, 103851. https://doi.org/10.1016/j.advwatres.2021.103851

Tapiador, F.J., Kidd, C., Hsu, K. -L., Marzano, F., 2004. Neural networks in satellite rainfall estimation. Meteorological Applications 11, 83–91. https://doi.org/10.1017/S1350482704001173

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An Overview of CMIP5 and the Experiment Design. Bulletin of the American Meteorological Society 93, 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Tefera, M.L., Giovanna Seddaiu, Alberto Carletti, Awada, H., 2025. Rainfall variability and drought in West Africa: challenges and implications for rainfed agriculture. Theor Appl Climatol 156, 41. https://doi.org/10.1007/s00704-024-05251-8

Teng, J., Croke, B., Tan, D., Iwanaga, T., Jakeman, A.J., Pollino, C., Stratford, D., Vaze, J., Dawes, W., Bridgeman, P., Sengupta, A., 2023. Development of a computationally efficient floodplain ecological response model for large-scale, data-sparse riparian environments. Ecological Informatics 77, 102252. https://doi.org/10.1016/j.ecoinf.2023.102252

Tikhamarine, Y., Souag-Gamane, D., Ahmed, A.N., Sammen, S.Sh., Kisi, O., Huang, Y.F., El-Shafie, A., 2020. Rainfall-runoff modelling using improved machine learning methods: Harris hawks optimizer vs. particle swarm optimization. Journal of Hydrology 589, 125133. https://doi.org/10.1016/j.jhydrol.2020.125133

Toure, A., Diekkrüger, B., Mariko, A., 2016. Impact of Climate Change on Groundwater Resources in the Klela Basin, Southern Mali. Hydrology 3, 17. https://doi.org/10.3390/hydrology3020017

Touré, M., Sanogo, D., Dembele, S., Diawara, S.I., Oppfeldt, K., Schiøler, K.L., Haidara, D.B., Traoré, S.F., Alifrangis, M., Konradsen, F., Doumbia, S., 2016. Seasonality and shift in age-specific malaria prevalence and incidence in Binko and Carrière villages close to the lake in Selingué, Mali. Malar J 15, 219. https://doi.org/10.1186/s12936-016-1251-4

Traore, S.S., 2015. Long -Term Vegetation Dynamics over the Bani River Basin as Impacted by Climate Change and Land Use (Thesis). Kwame Nkrummah University of Science and Technology.

Vreysen, M.J.B., Balenghien, T., Saleh, K.M., Maiga, S., Koudougou, Z., Cecchi, G., Bouyer, J., 2013. Release-Recapture Studies Confirm Dispersal of Glossina palpalis gambiensis between River Basins in Mali. PLoS Negl Trop Dis 7, e2022. https://doi.org/10.1371/journal.pntd.0002022

Ward, P.J., Jongman, B., Kummu, M., Dettinger, M.D., Sperna Weiland, F.C., Winsemius, H.C., 2014. Strong influence of El Niño Southern Oscillation on flood risk around the world. Proc. Natl. Acad. Sci. U.S.A. 111, 15659–15664. https://doi.org/10.1073/pnas.1409822111

Wu, Z., Wang, S., Zhao, J., Chen, L., Meng, H., 2014. Synergistic effect on thermal behavior during co-pyrolysis of lignocellulosic biomass model components blend with bituminous coal. Bioresource Technology 169, 220–228. https://doi.org/10.1016/j.biortech.2014.06.105

Xiang, H., Xi, Y., Mao, D., Mahdianpari, M., Zhang, J., Wang, M., Jia, M., Yu, F., Wang, Z., 2023. Mapping potential wetlands by a new framework method using random forest algorithm and big Earth data: A case study in China's Yangtze River Basin. Global Ecology and Conservation 42, e02397. https://doi.org/10.1016/j.gecco.2023.e02397

Xiang, X., Li, Q., Khan, S., Khalaf, O.I., 2021. Urban water resource management for sustainable environment planning using artificial intelligence techniques. Environmental Impact Assessment Review 86, 106515. https://doi.org/10.1016/j.eiar.2020.106515

Xu, R., Chen, N., Chen, Y., Chen, Z., 2020. Downscaling and Projection of Multi-CMIP5 Precipitation Using Machine Learning Methods in the Upper Han River Basin. Advances in Meteorology 2020, 1–17. https://doi.org/10.1155/2020/8680436

Yaseen, Z.M., Sulaiman, S.O., Deo, R.C., Chau, K.-W., 2019. An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. Journal of Hydrology 569, 387–408. https://doi.org/10.1016/j.jhydrol.2018.11.069

Yue, S., Wang, C.Y., 2002. Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test. Water Resources Research 38. https://doi.org/10.1029/2001WR000861

Annex

List of publications

Published (Framework of the PhD)

Kouyaté, F.; Guédjé, F.K.; Ndiaye, A.; Ganni Mampo, O.M. (2025). Spatial and Temporal Variability of Extreme Hydroclimatic Events in the Bani River Basin. Hydrology, 12(5).

https://doi.org/10.3390/hydrology12010005

Kouyaté, F.; Aminzadeh, M.; Guédjé, F.K.; Madani, K.; Shokri, N. (2024). Utilizing Satellite Data and Machine Learning Algorithms to Predict Water Level Fluctuations in the Bani River in Africa. Poster Presentation American Geophysical Union (AGU) Fall Meeting, Washington D.C., USA.

https://doi.org/10.13140/RG.2.2.12200.48649

Under Review (Framework of the PhD)

Kouyaté, F.; Aminzadeh, M.; Guédjé, F.K.; Madani, K.; Shokri, N. (Under review). Predictive Modeling of River Water Level Fluctuations: A Case Study of the Bani River in Africa. Journal of Hydrology: Regional Studies. Journal of Hydrology: Regional Studies (Elsevier)

Kouyaté, F.; Guédjé, F.K.; Fofana, M. (Accepted with revisions). Climate Projections for Assessing Hydroclimatic Extremes in the Bani River Basin, West Africa. Journal of Environmental Challenges (Elsevier).

Others

Bonkoungou, B., Bossa, A. Y., van der Kwast, J., Mul, M., Kassambara, B., Kouyaté, F., Yameogo, T. S., Ahoton, D. H., & Sintondji, L. O. (Under Review). Water accounting in the Inner Niger Delta under climate and land use changes. Hydrology Research Journal (IWA)

Workshop to Validate the Floodplain Atlas of Bamako (June 21-22, 2023): Member, GIZ-Niger Basin Authority.

WASCAL Scientific Symposium 2022 (December 6-9, 2022): Oral Presentation, held in Ouagadougou, Burkina Faso.

Conference “Intermittency in Headwater Streams - Challenges and Opportunities from an Interdisciplinary Perspective” (June 4-6, 2024): Poster Presentation, held at the University of Bonn, Germany.

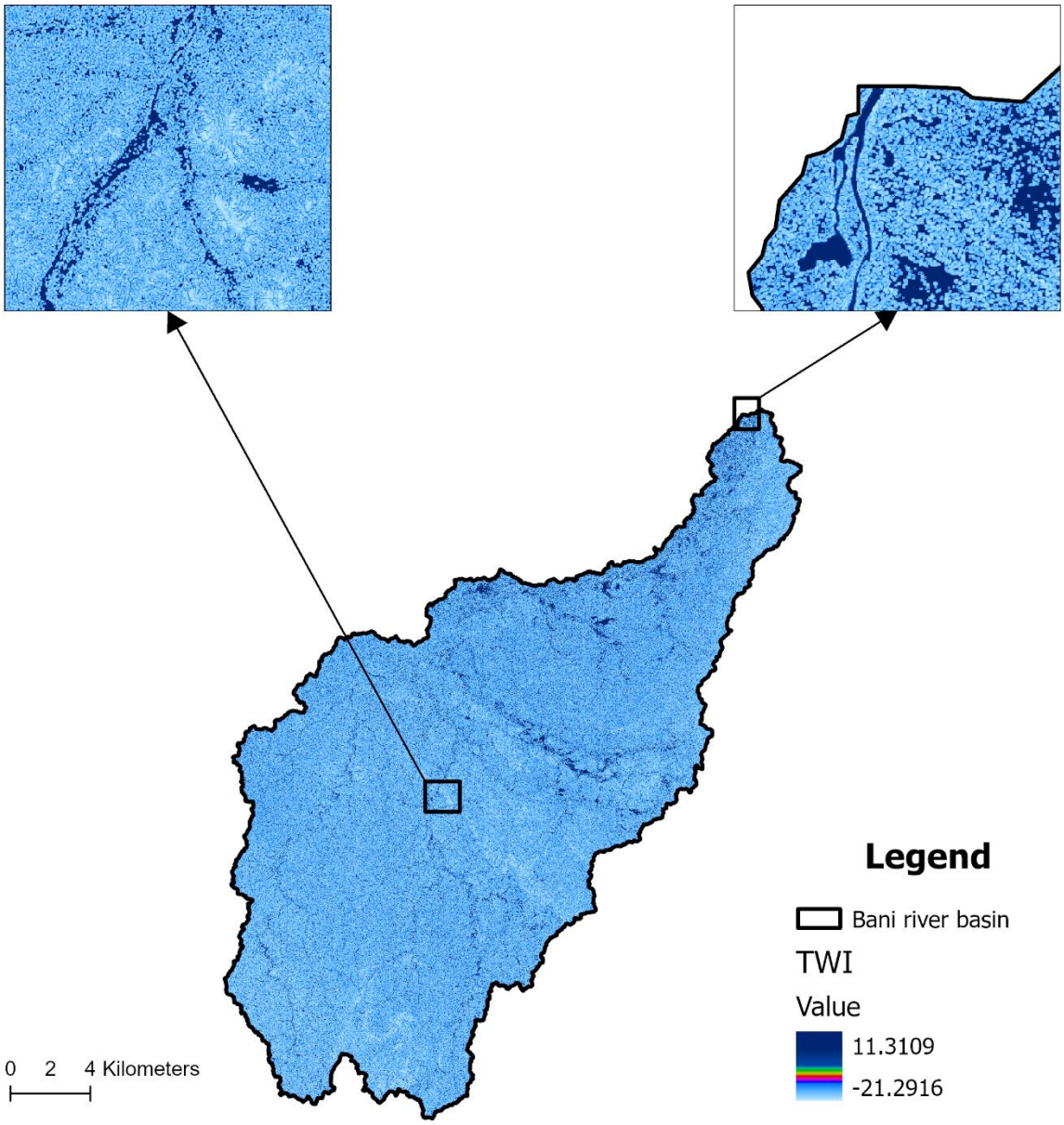

Figure 30: Topographic Wetness Index (TWI) of the BRB

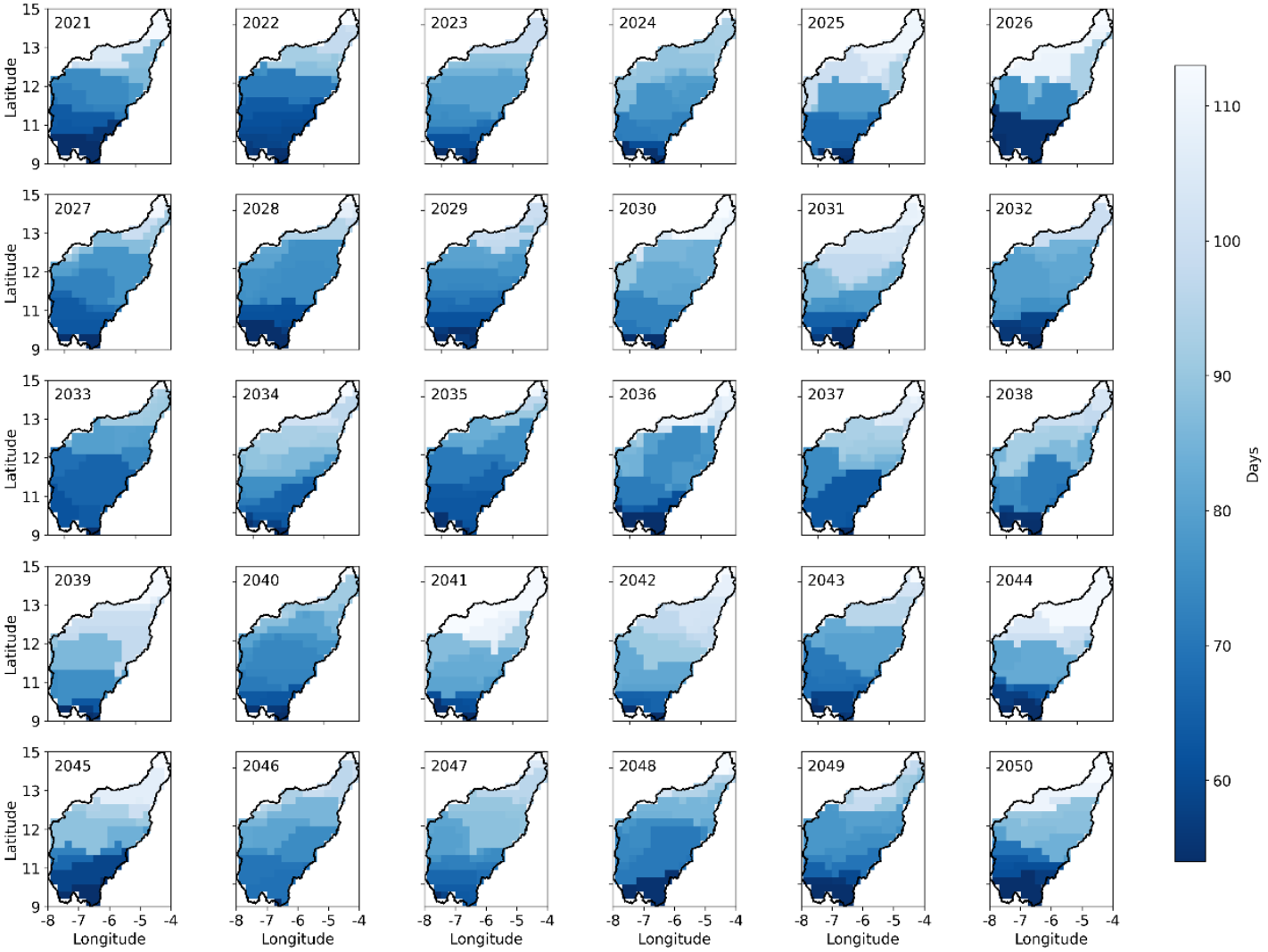

Figure 31: Spatial distribution of CDD in the BRB from 2021 to 2050 under the SSP5-8.5 scenario.

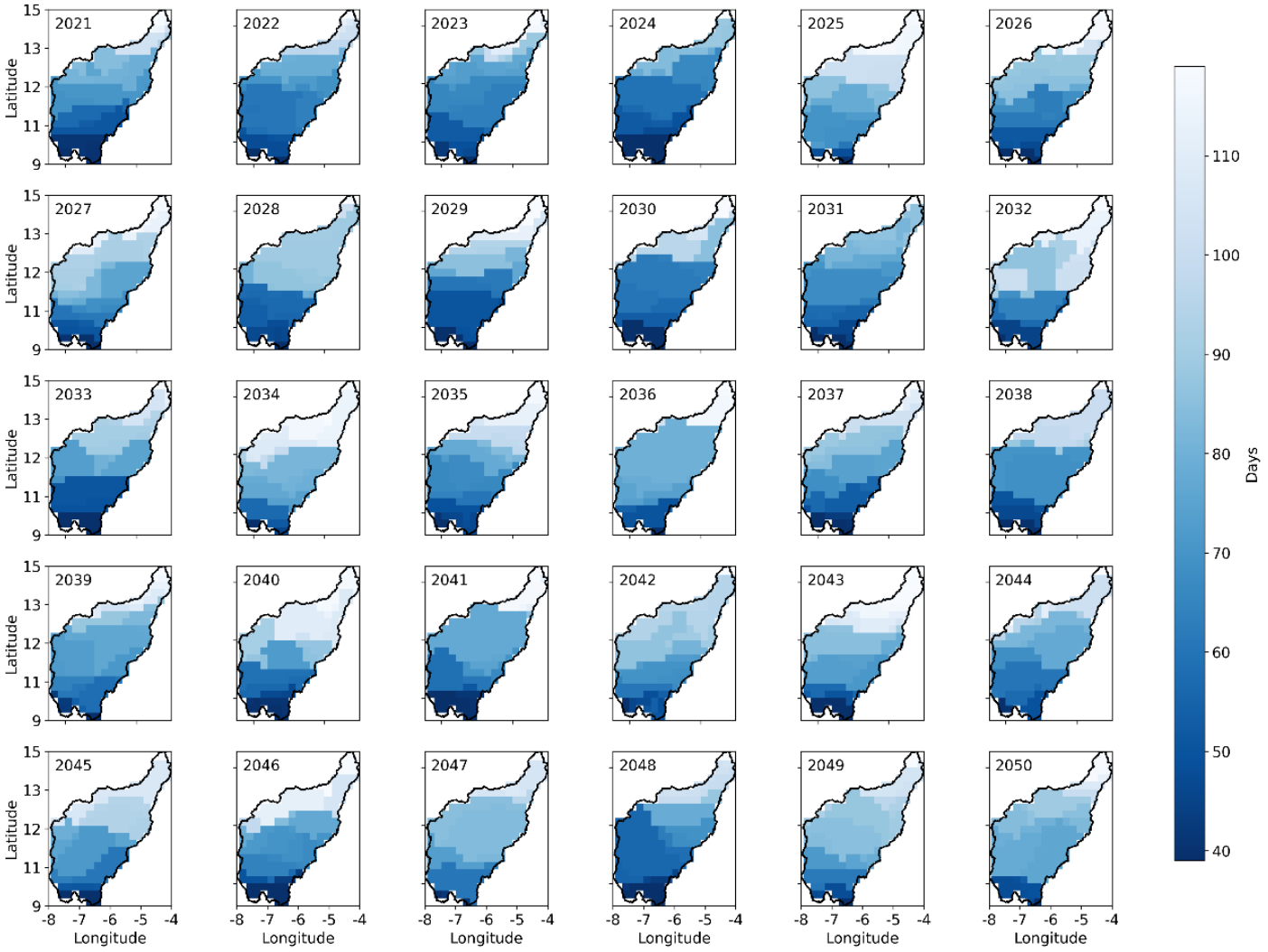

Figure 32: Spatial distribution of CDD in the BRB from 2021 to 2050 under the SSP2-4.5 scenario.

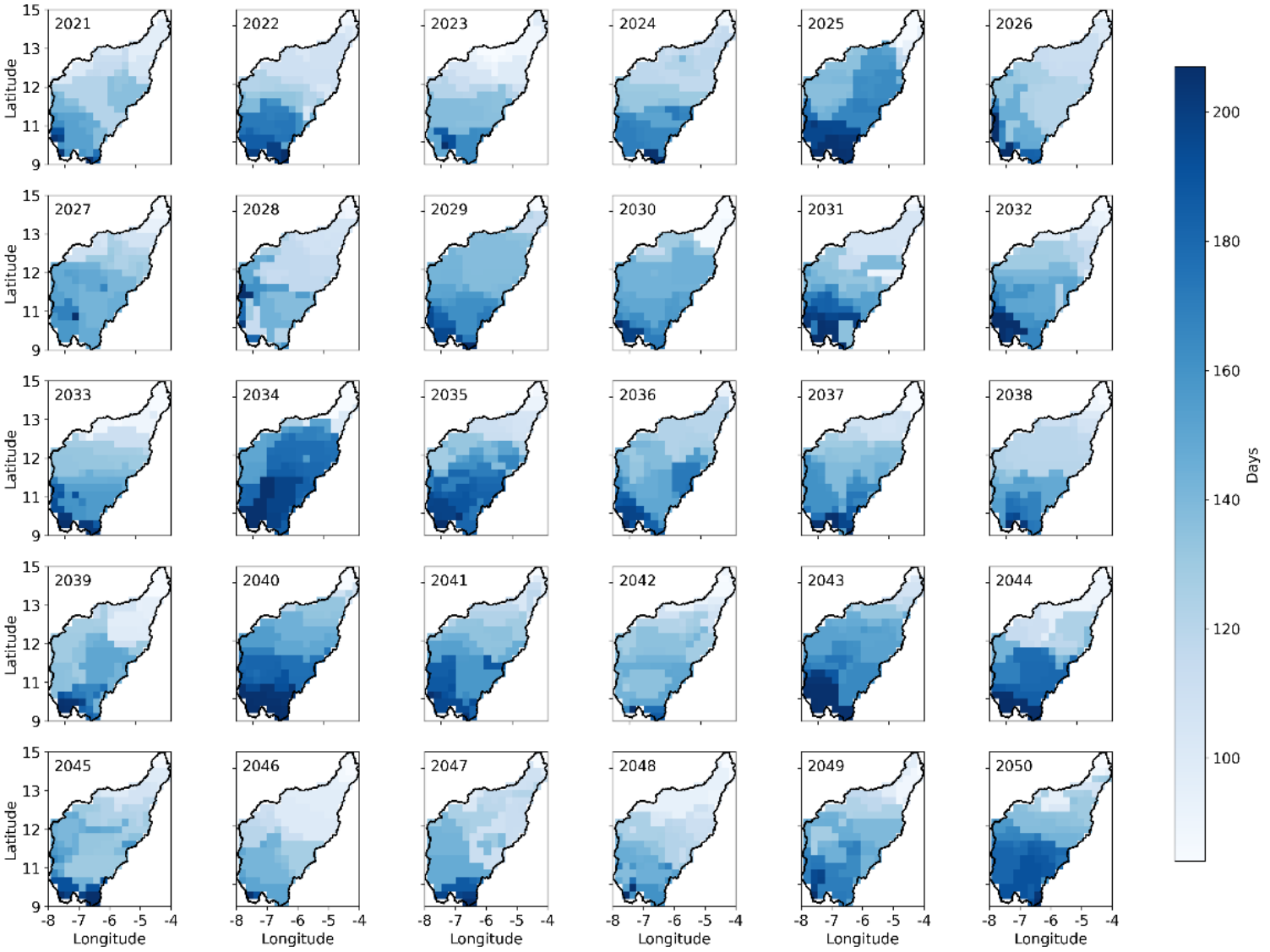

Figure 33: Spatial distribution of CWD in the BRB from 2021 to 2050 under the SSP5-8.5 scenario.

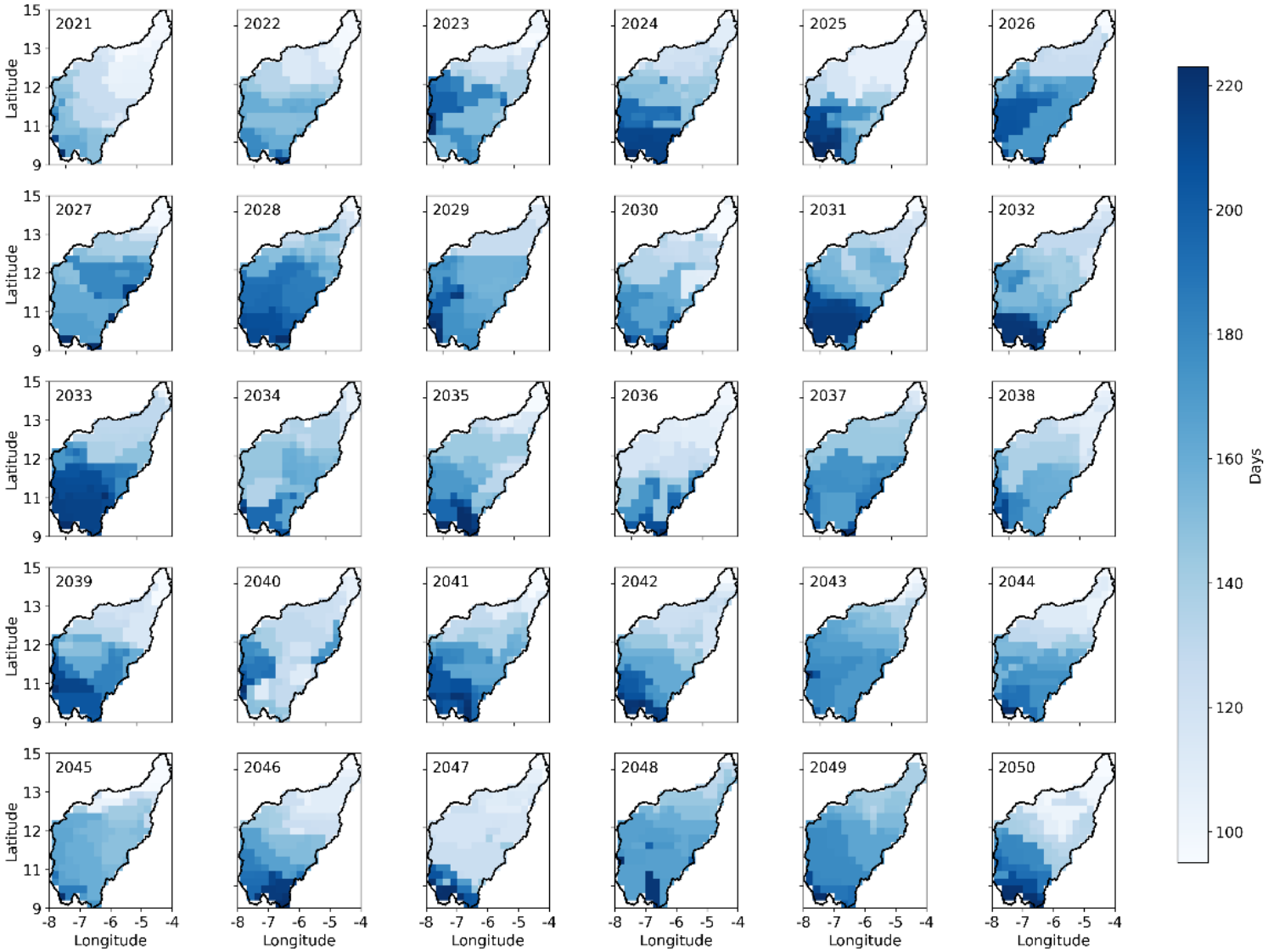

Figure 34: Spatial distribution of CWD in the BRB from 2021 to 2050 under the SSP2-4.5 scenario.

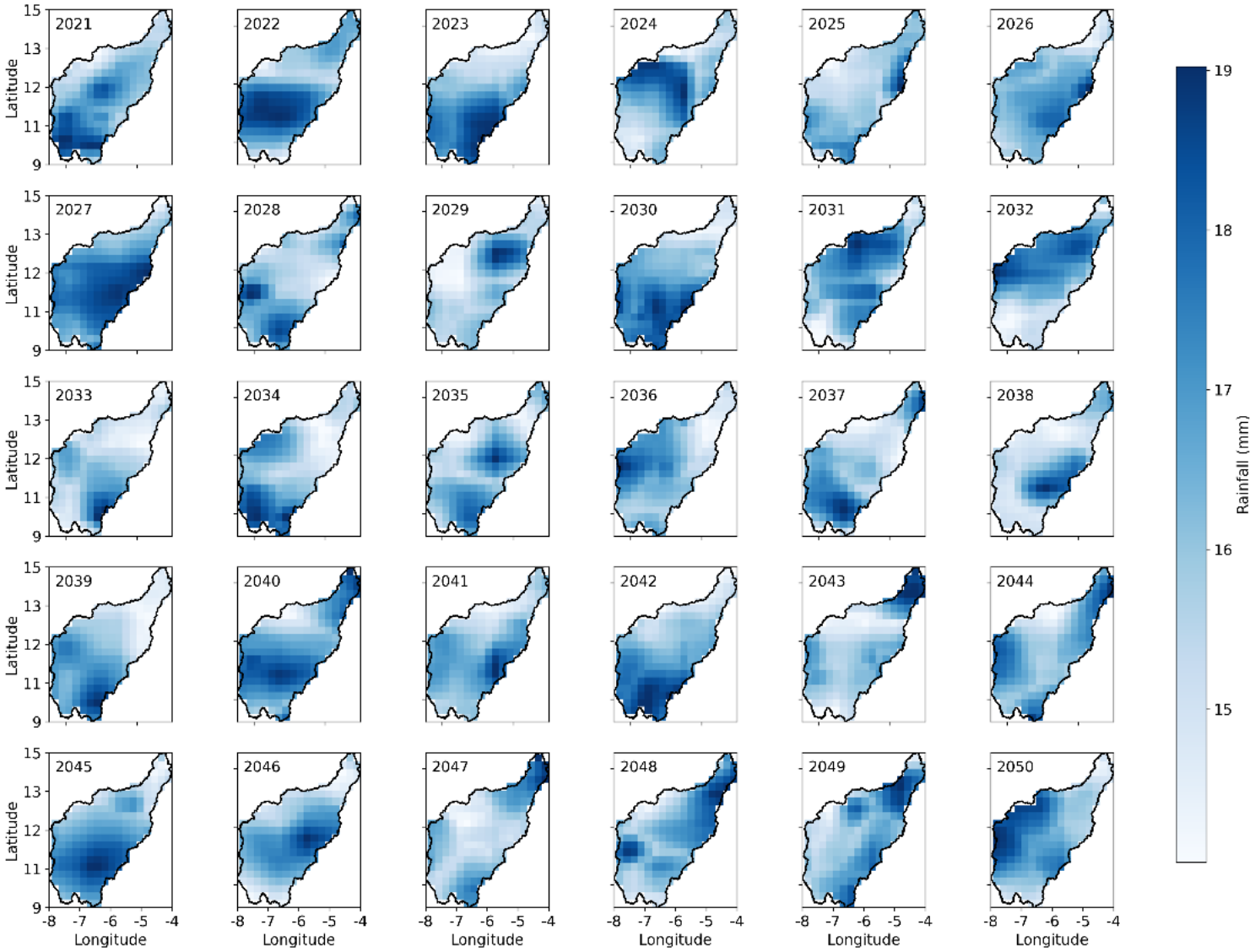

Figure 35: Spatial distribution of RX1DAY in the BRB from 2021 to 2050 under the SSP5-8.5 scenario.

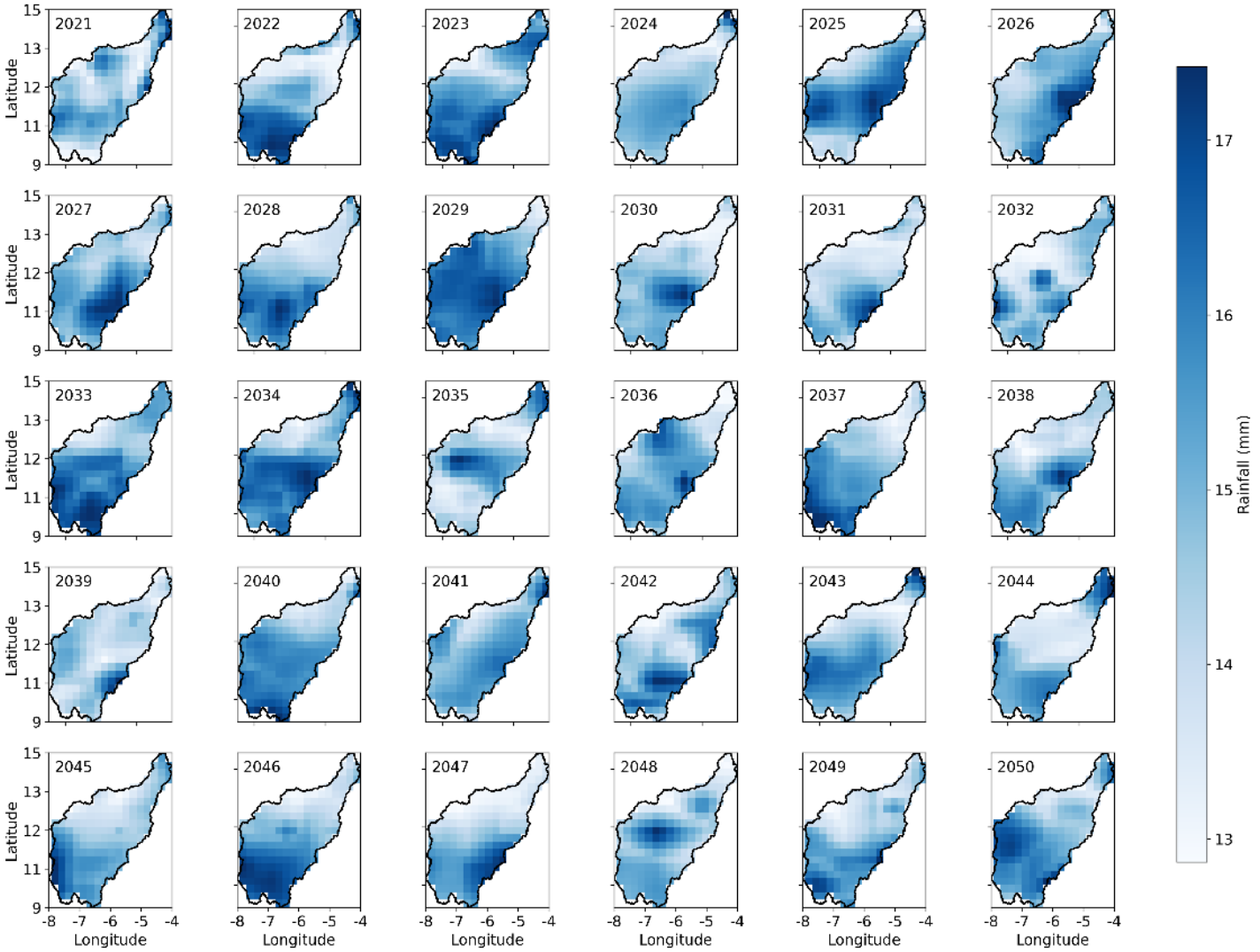

Figure 36: Spatial distribution of RX1DAY in the BRB from 2021 to 2050 under the SSP2-4.5 scenario.

I am Fousseini KOUYATE, born on November 30, 1996, in Sikasso, Mali. I earned my Bachelor's degree in Geology, specializing in Geo-resources, from Ibn Tofail University in Morocco in 2018, followed by a Master's degree in Hydroinformatics and Hydrosystems Management from the same institution in 2021. My early academic work focused on hydrological and hydrogeological modeling of water resources. In 2020, I completed a research placement at the Ministry of Water and Sanitation of Senegal, where I analyzed the hydrological performance of satellite-based precipitation products in the Bani River Basin. After my master's defense, I was awarded the WASCAL Ph.D. scholarship in Climate Change and Water Resources, funded by the German Federal Ministry of Education and Research (BMBF), at the University of Abomey-Calavi, Benin.

Since March 2022, I have been affiliated with the Institute of Geo-Hydroinformatics at Hamburg University of Technology (TUHH). My Ph.D. research focuses on applying artificial intelligence to predict river water-level fluctuations and to support climate-resilient water management in West Africa.

As part of my doctoral work, I have also contributed to several World Bank–funded initiatives in Mali, including projects aimed at strengthening early warning systems, improving flood forecasting, and supporting landscape restoration and water-security planning.

Abstract:

This study addresses the hydroclimatic challenges of the Bani River Basin (BRB) in southeastern Mali, where data scarcity, climate variability, and socio-political instability severely complicate water resource management. First, we validate satellite-based rainfall products (e.g., CHIRPS) against available ground observations to determine their reliability in capturing key precipitation patterns. The results confirm the usefulness of these products, particularly in regions where insecurity and logistical constraints limit in-situ measurements. Second, we employ statistical trend analyses to examine the spatiotemporal evolution of rainfall and river discharge indices over recent decades. The findings reveal shifting extremes more frequent short-duration heavy rainfall and intensifying peak flows, raises concerns about water scarcity, flood risks, and agricultural stability. Third, we develop and test machine learning (ML) models, including Random Forest and Gradient Boosting, to replicate historical water-level fluctuations. These models demonstrate strong predictive capability, outperforming traditional methods when faced with complex, nonlinear datasets. Finally, by integrating Coupled Model Intercomparison Project Phase 6 (CMIP6) projections under scenarios SSP2-4.5 and SSP5-8.5, we forecast future water levels and identify key breakpoints in hydrological regimes. The results underscore an urgent need for adaptive management strategies, including improved monitoring infrastructures, data-sharing mechanisms, and community-driven interventions. Overall, this research offers actionable insights for policymakers and stakeholders, emphasizing the viability of satellite-based data and advanced computational methods to mitigate climate risks and strengthen water governance in the BRB.

Key words: Bani River Basin; Data scarcity; Socio-political instability; Machine learning; Satellite-based rainfall; CMIP6 scenarios