UNIVERSITE JOSEPH KI-ZERBO

------------

ECOLE DOCTORALE INFORMATIQUE ET CHANGEMENT CLIMATIQUE (EDICC)

BURKINA FASO

*La Patrie ou la Mort, nous Vaincrons*

Order N° : E00949120161

## MASTER RESEARCH PROGRAM

## SPECIALITY: INFORMATICS FOR CLIMATE CHANGE (ICC)

# MASTER THESIS

## Subject:

# Statistical Downscaling of Global Climate Models for Temperature and Precipitation over West Africa Using Machine Learning Algorithms

**Presented 07th July 2025, by:**

**ADJA Kodjo Aménoagbé**

**Examination jury**

**President: Pr DAHO Tizane**

**Members :**

- **Dr Zongo Augustin** (external examiner)
- **Dr Yaya TRAORE** (supervisor);
- **Dr Windmanagda SAWADOGO** (co-supervisor)

**Academic year 2024-2025**

# DEDICATION

To my beloved Mother, my dear Father, and my precious Sisters with deepest gratitude for your unwavering love, sacrifice, and prayers.

***"May the Lord bless you and keep you;***

***may His face shine upon you and be gracious to you."***

# ACKNOWLEDGEMENTS

All praise be to our Lord Jesus Christ, the Author of life and Sustainer of all creation. By His grace, we have been given strength, wisdom, and endurance throughout this academic journey. We give glory to God for breathing life into every step of this endeavor and for granting us the perseverance and discipline required to complete this thesis.With heartfelt gratitude, we also extend sincere appreciation to:

- German Ministry of Education and Research (BMBF) for funding this program;
- Pr. Amadé Ouédraogo , Director of the "École Doctorale Informatique pour le Changement Climatique" (ED-ICC) for his faith in this institute and for all his efforts;
- Dr. Ousmane COULIBALY, Deputy Director of ED-ICC for listening to the needs of students and for having transmitted to us the meaning of a work well done and commensurate with the training obtained;
- West African Science Service Centre on Climate Change and Adapted Land Use (WASCAL) for providing this Graduate Study Program (GSP) of capacity building and scholarship;
- Prof. Yaya TRAORE, my Major Supervisor who despite his numerous occupations supervised and framed our thesis work;
- Dr. Windmanagda SAWADOGO, my Co-Supervisor for showing interest since the early stages of this work;
- Dr. Benewindé Jean-Bosco ZOUNGRANA, Scientific Coordinator of ED-ICC for sharing his follow-up and advice, and through him all ED-ICC staff members;
- Mr N'KOYI M'poh, my Internship Supervisor, for showing great interest in my work and providing me with all the required support during my internship at ANAMET;
- Dr. Latifou ISSAOU the General Director of Togo National Agency of Meteorology (ANAMET), for his permission for the internship;
- Prof. Komi AGBOKA, the Director of WASCAL Togo, for being so welcoming and finding me accommodation at "Université de Lomé" during my period of internship;
- All my colleagues and seniors (special thanks to my senior BOUBAKAR from WASCAL Batch 2), for their unconditional support and confidence;
- All those who, from near or far, have contributed to the drafting of this document.

# ABSTRACT

Global Climate Models (GCMs) provide essential projections for understanding future climate conditions but operate at coarse spatial resolutions, limiting their applicability for regional and local-scale climate impact assessments. To address this gap, this study applies advanced machine learning (ML) based statistical downscaling techniques to improve the spatial resolution and accuracy of temperature and precipitation projections over West Africa. Using historical climate observations (1985–2014) and future CMIP6 GCM outputs (2015–2100), we implement and evaluate both regression-based and deep learning (DL) models specifically, Convolutional Neural Networks (CNNs) and Convolutional Long Short-Term Memory networks (ConvLSTM).

The study assesses downscaling performance across multiple GCMs (CMCC-ESM2, CMCC-CM2-SR5, FGOALS-g3, NorESM2-LM), comparing model outputs against high-resolution reference datasets (ERA5 for temperature and CHIRPS for precipitation). Results reveal that CNN models are particularly effective in downscaling precipitation, offering high spatial fidelity, reduced biases, and superior accuracy (RMSE = 0.137 mm/day, $R^2$ = 0.995), especially in humid and orographic zones. Conversely, ConvLSTM models outperform CNNs in temperature downscaling, capturing temporal dependencies and seasonality with high precision (RMSE = 0.212 °C, $R^2$ = 0.990, r = 0.996). While both DL models significantly outperform raw GCM outputs, CNN is better suited for applications requiring spatial detail, such as rainfall intensity mapping, while LSTM is more appropriate for time-sensitive applications, including temperature forecasting and hydrological modeling.

This research confirms the potential of deep learning–based statistical downscaling to enhance the usability of GCM outputs for regional climate analysis and adaptation planning in West Africa. The results contribute to ongoing efforts to build robust, high-resolution climate datasets that can support informed policy decisions and sectoral planning under climate change.

**Keywords:** Statistical downscaling, global climate models, machine learning, West Africa, climate projections.

# RESUMÉ

Les modèles climatiques globaux (GCMs) fournissent des projections essentielles pour comprendre les conditions climatiques futures, mais leur résolution spatiale grossière limite leur applicabilité pour les évaluations d'impacts climatiques à l'échelle régionale et locale. Pour combler cette lacune, cette étude applique des techniques avancées de réduction statistique d'échelle basées sur l'apprentissage automatique (machine learning - ML) afin d'améliorer la résolution spatiale et la précision des projections de température et de précipitations en Afrique de l'Ouest.

En utilisant des observations climatiques historiques (1985–2014) ainsi que des projections issues des GCMs CMIP6 pour la période future (2015–2100), l'étude met en œuvre et évalue à la fois des modèles de régression et des modèles d'apprentissage profond (deep learning – DL), en particulier les réseaux de neurones convolutifs (CNN) et les réseaux de mémoire à long terme convolutifs (ConvLSTM).

La performance des modèles de réduction d'échelle est évaluée à partir de plusieurs GCMs (CMCC-ESM2, CMCC-CM2-SR5, FGOALS-g3, NorESM2-LM) en les comparant à des jeux de données de référence à haute résolution (ERA5 pour la température et CHIRPS pour les précipitations). Les résultats montrent que les modèles CNN sont particulièrement efficaces pour la réduction d'échelle des précipitations, offrant une fidélité spatiale élevée, des biais réduits et une excellente précision (RMSE = 0,137 mm/jour, $R^2$ = 0,995), notamment dans les zones humides et orographiques. À l'inverse, les modèles ConvLSTM surpassent les CNN pour la réduction d'échelle de la température, en capturant avec précision les dépendances temporelles et la saisonnalité (RMSE = 0,212 °C, $R^2$ = 0,990, r = 0,996).

Bien que les deux modèles DL améliorent considérablement les performances par rapport aux GCMs bruts, le CNN est mieux adapté aux applications nécessitant une précision spatiale, comme la cartographie de l'intensité des précipitations, tandis que le LSTM est plus pertinent pour les applications sensibles au temps, telles que la prévision de température ou la modélisation hydrologique.

Cette recherche confirme le potentiel de la réduction statistique d'échelle basée sur l'apprentissage profond pour améliorer la pertinence des projections issues des GCMs en vue d'analyses climatiques régionales et de la planification de l'adaptation en Afrique de l'Ouest. Les résultats contribuent aux efforts actuels visant à développer des bases de données climatiques robustes et à haute résolution pour soutenir la prise de décision éclairée et la planification sectorielle face au changement climatique.

**Mots-clés** Désagrégation statistique, modèles climatiques globaux, apprentissage automatique, Afrique de l'Ouest, prévisions climatiques.

# Table of Contents

## List of tables



## List of figures

# ACRONYMS AND ABBREVIATIONS

| | |
|---|---|
| **CHIRPS** | Climate Hazards Group InfraRed Precipitation with Station data |
| **CNN** | Convolutional Neural Network |
| **GCM** | Global Climate Model |
| **LSTM** | Long Short-Term Memory |
| **P95** | 95th Percentile Precipitation |
| **Anomaly** | Deviation from the mean precipitation |
| **mm/day** | Millimeters per day, unit of precipitation |
| **Downscaled** | High-resolution output derived from coarse GCM data |
| **Bias** | Difference between modeled and observed values |
| **Near** | Near-term projection (2015 - 2060) |
| **Far** | Far-term projection (2061 - 2100) |
| **CMCC-CM2-SR5** | Centro Euro-Mediterraneo sui Cambiamenti Climatici Climate Model (CMIP6 generation) |
| **NorESM2-LM** | Norwegian Earth System Model, Low Resolution |
| **FGOALS-g3** | Flexible Global Ocean-Atmosphere-Land System Model, Grid-point Version 3 |
| **ConvLSTM** | Convolutional Long Short-Term Memory |
| **ERA5** | ECMWF Reanalysis v5, a high-resolution climate reanalysis dataset |
| **ML** | Machine Learning |
| **HPC** | High Performance Computing |
| **WAM** | West African Monsoon |

# INTRODUCTION

## I. Background

Climate change is one of the most serious and urgent challenges of the 21st century, with far-reaching environmental, economic, and social impacts. One of the most vulnerable regions to its impacts is West Africa, a climatic hotspot where interannual variability is high and the region is especially sensitive to increases in global warming. The region is sensitive to various mechanisms such as the West African Monsoon (WAM), land-atmosphere feedbacks, and large-scale ocean-atmosphere teleconnections, among others related to the Atlantic Multidecadal Oscillation and ENSO (Janicot et al. 2021). This has major implications for agriculture, water resources, human health, and disaster risk across the region. In this context, climate information will be vital for adaptation, early warning, and sustainable development planning.

West Africa is highly vulnerable to climate change due to its dependence on rain-fed agriculture, water resources, and natural ecosystems (Lebel et al., 2003). Extreme climatic events such as droughts and floods have increased in frequency, impacting food security and livelihoods (IPCC, 2021). Additionally, rapid urbanization and land-use changes exacerbate environmental degradation, further increasing the region's climate vulnerability.

Nonetheless, Global Climate Models (GCMs), which are the main tools used to simulate future climate scenarios, are inherently limited in their representation of regional-scale processes. GCMs have a too coarse resolution of 100-250 km, and are therefore not able to simulate local climate variability and topographical influences, which are particularly relevant in West Africa, where mesoscale features such as coastal rainfall gradients, rainfall orographic enhancement, and convection processes are all important (Benestad et al., 2017; Sun et al., 2021).

Statistical downscaling methods have been developed to complement dynamical methods to overcome this spatial mismatch. Statistical downscaling uses historical relationships between large-scale atmospheric conditions (the predictors) and local climate conditions (the predictands) to generate projections of climate at fine resolution based on coarse outputs from GCMs (Gaitan et al., 2014).

New developments in machine learning (ML) have greatly improved the performance and flexibility of statistical downscaling models. ML methods, including Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Random Forests (RFs), and more recently deep learning architectures such as Convolutional Neural Networks (CNNs) and Convolutional Long Short-Term Memory networks (ConvLSTM), represent powerful means to capture the complex, nonlinear, and often non-stationary relationships between predictors and predictands (Maraun et al., 2010; Vandal et al., 2019; Baño-Medina et al., 2021).

These methods have shown great promise in improving the spatial accuracy, temporal resolution, and physical coherence of downscaled climate data, particularly for key variables such as temperature and precipitation. Given the lack of observational data available for West Africa in particular, and the fact that millions of people will be affected by climate extremes globally, high-resolution ML downscaling represents a promising avenue to generate robust and useful climate information.

This research expands on these advances to assess the skill of CNN and ConvLSTM models to statistically downscale temperature and precipitation projections from four different CMIP6 GCMs, CMCC-CM2-SR5, FGOALS-g3, CMCC-ESM2, and NorESM2-LM, over West Africa. The analysis attempts to evaluate the accuracy, bias correction, and spatial improvement of ML downscaled data using high-resolution datasets for observations, including CHIRPS and ERA5.

## II. Problem Statement

Climate models are indispensable tools for projecting future climate conditions, yet their spatial resolution and systematic biases remain critical challenges, particularly in vulnerable regions such as West Africa. Global Climate Models (GCMs), the backbone of long-term climate projections, typically operate at coarse spatial resolutions (100–250 km), which are insufficient for capturing the fine-scale atmospheric processes and regional climatic heterogeneity driven by the West African Monsoon system (Sylla et al., 2013).

Several studies have documented the systematic biases in GCM outputs over West Africa, especially regarding precipitation. GCMs often: (1) Underestimate rainfall in monsoon zones, such as coastal Nigeria and Ghana. (2) Overestimate precipitation in dry inland zones, leading to misrepresentation of hydrological extremes. (3) Fail to resolve mesoscale convective systems, which are essential to precipitation variability in the Sahel (Dosio et al., 2015; Nikulin et al., 2012).

One of the main issues with GCMs, which has not been resolved despite the great strides made in climate modelling, is that the models' relatively coarse scale makes the results difficult to use for regional climate impact assessment (Wilby & Dawson, 2013). Particularly in West Africa, where geography is varied and climate is complex, climate data at a higher resolution is vital for effective policy making. Significantly, conventional statistical downscaling techniques, including regression-based techniques and weather generators are typically not able to represent the complex nature of the relationships between large-scale predictors and local climate responses (Zorita & von Storch, 1999). Although successful, the conventional dynamical downscaling techniques are computationally complex and often lengthy (Quenum, et al, 2019).

As highlighted by Omaya and Muthama (2024), previous downscaling initiatives in West Africa have faced persistent challenges, including poor data quality, limited computational capacity, and the absence of robust models capable of generalizing across the region's heterogeneous climate zones. They argue that addressing these issues requires the integration of advanced machine learning (ML) techniques, which can generate more accurate and spatially disaggregated climate information at actionable scales. ML approaches offer the advantage of improved predictive performance while reducing dependence on costly High-Performance Computing (HPC) infrastructure, making them particularly suitable for regions with constrained resources.

Comprehending and forecasting temperature and precipitation trends is crucial to enhance regional climate modeling and, as a result, bolster adaptation strategies in West Africa. This study seeks to use statistical downscaling approaches to develop machine learning models that produce optimal machine learning output refinements of temperature and precipitation from GCMs, even in low-resource contexts. The use of advanced machine learning algorithms will improve the accuracy of climate projections and ultimately allow for better climate adaptation and mitigation practices in the region.

## III. Research Questions

Accurate and spatially refined climate projections are essential for effective adaptation planning in West Africa, a region characterized by high vulnerability to climate variability. Machine learning (ML) techniques offer promising alternatives to conventional approaches for downscaling climate data, particularly in correcting biases and enhancing the spatial resolution of temperature and precipitation outputs from coarse-resolution Global Climate Models (GCMs).

To guide this investigation, the following research questions are posed:

- **Q1.** Which machine learning models are most effective at enhancing the spatial resolution and reducing bias in temperature projections over West Africa?
- **Q2.** How do different ML-based downscaling models perform in improving the accuracy of precipitation projections derived from GCMs in the region?

## IV. Research hypotheses

To address the research questions, the following hypotheses are underpinned:

- **H1.** Deep learning models, particularly Convolutional Neural Networks (CNNs), are more effective than ConvLSTM and traditional ML techniques in improving the spatial resolution of downscaled temperature data over West Africa, relative to raw GCM outputs.
- **H2.** Advanced ML models, including ConvLSTM architectures, provide more accurate and reliable downscaling of precipitation data compared to raw GCM outputs, resulting in better representation of regional climatic patterns across West Africa.

## V. Research objectives

To evaluate machine learning-based statistical downscaling techniques for improving the accuracy and spatial resolution of temperature and precipitation projections over West Africa

O1. To compare the effectiveness of various ML models, including deep learning techniques like CNNs and ConvLSTM, in downscaling temperature data for West Africa with the raw GCMs.

O2. To evaluate the performance of different ML models, particularly deep learning approaches, in downscaling precipitation data for improving the accuracy of precipitation projections over West Africa with the raw GCMs.

## VI. Thesis structure

This thesis is structured to offer a comprehensive assessment of machine learning–based statistical downscaling of climate projections over West Africa. It begins with an introduction that outlines the general context and rationale of the study, followed by the problem statement, research questions, hypotheses, and specific objectives. The first chapter presents a review of relevant literature, covering statistical downscaling methods, the limitations of Global Climate Models (GCMs), and the growing application of machine learning techniques in climate modeling. Chapter Two describes the study area and the methodology, including observational climate records and outputs from selected CMIP6 GCMs for temperature and precipitation downscaling. Chapter Three focuses on the results and discussion, presenting a detailed evaluation of the performance of Convolutional Neural Networks (CNN) and Convolutional Long Short-Term Memory (ConvLSTM) models. It addresses the first research question by comparing downscaled outputs against observed datasets using both statistical metrics and spatial analyses. And also provides a critical discussion of the findings, examining the strengths and limitations of the ML models, their implications for climate risk management, and addressing the second research question. The final section, Conclusion and Perspectives, summarizes the main outcomes of the study, reflects on the research objectives, and outlines future directions for improving statistical downscaling approaches and their applications in climate adaptation planning.

# CHAPTER 1: LITERATURE REVIEW

## 1.1 Introduction

Climate change is one of the most serious and pressing global issues, especially relevant for highly climate-sensitive regions and sectors such as agriculture and water in places like West Africa. Reliable climate change projections are required for the development of sound adaptation and mitigation strategies. But, Global Climate Models (GCMs), which model climate at a global scale, are set at coarse spatial resolutions (250-600 km) and therefore are not adequate for regional or local applications. This scale misfit emphasizes the need for downscaling methods that can provide estimates transferred from general climate models to specific localities.

## 1.2 Downscaling Techniques

Downscaling encompasses a set of methods that are used to obtain local-to-regional-scale climate data from GCM output at a coarse resolution. These methods are crucial in impact assessments, vulnerability mapping, and climate-aware planning, all of which can be critically important in data-poor regions such as West Africa. Downscaling methodologies may be broadly categorized into dynamical and statistical downscaling, each having its own advantages and disadvantages.

### 1.2.1 Dynamical Downscaling

Dynamical downscaling exploits the nesting of high-resolution Regional Climate Models (RCMs) within GCMs to represent climate processes at a finer scale. These RCMs simulate local features such as orographic precipitation, coastal winds, and convective processes through complex atmospheric, oceanic, and land-surface interactions. The Coordinated Regional Climate Downscaling Experiment (CORDEX), and in particular the CORDEX-Africa framework, has allowed for the production of RCM-based projections over the African continent. A few examples include the work of Nikulin et al. (2012), who showed how RCMs can be used to depict the spatial variability of rainfall and temperature over sub-Saharan Africa under different climate scenarios.
Similarly, Giorgi et al. (2009) found that RCMs offer improved simulation of monsoon systems and of extremes over complex topography areas, such as West Africa. But, dynamical methods are rather robust but very costly computationally, as they involve large input data and long computational time. They also tend to be biased, inheriting the biases of their driving GCMs, and

also inter-model differences due to different physics schemes, boundary conditions, or parameterization techniques may occur (Rummukainen, 2010). But, dynamical downscaling continues to be a fundamental approach in areas where physical realism is important.

### 1.2.2 Statistical Downscaling

Statistical downscaling offers a cost-effective alternative for enhancing climate model outputs by establishing empirical relationships between large-scale atmospheric variables (predictors) and local climate variables (predictands).

This approach assumes that these relationships remain stationary over time, an assumption that may not hold under changing climate conditions (Wilby et al., 2004). Statistical downscaling methods can generally be categorized into three groups. First, weather typing methods classify synoptic-scale weather patterns and link them to local climate responses, as demonstrated by Hewitson and Crane (2006). Second, regression-based approaches develop linear or nonlinear relationships between predictor variables (such as geopotential height and humidity) and local climate indicators. Common techniques in this category include Multiple Linear Regression (MLR), Artificial Neural Networks (ANN), and Support Vector Machines (SVM). For instance, Chu et al. (2010) applied MLR to downscale daily temperature and precipitation over East Asia, noting its simplicity and interpretability. However, traditional regression models often struggle to represent extreme events and complex nonlinear relationships. To address these limitations, more flexible methods like ANN (Hessami et al., 2008) and SVM (Tripathi et al., 2006) have gained popularity due to their ability to capture nonlinear interactions and adapt to diverse data structures. Finally, stochastic weather generators represent a third group, which simulate synthetic climate sequences that statistically resemble historical observations, as described by Semenov and Barrow (1997).

In particular, among these, regression based methods are the most widely applied for their simplicity and reasonably good accuracy. For example, Chu et al. 2010 used MLR to downscale daily mean temperature, pan evaporation, and precipitation and found its utility "an asset to existing climate scenarios."

## 1.3 Machine Learning in Statistical Downscaling

Machine Learning (ML) algorithms have gained increasing prominence in climate downscaling due to their capacity to model complex, nonlinear, and high-dimensional relationships between large-scale atmospheric predictors and local-scale climatic responses capabilities often beyond the reach of traditional statistical techniques. Unlike conventional regression models that typically assume linearity and stationarity, ML models can generalize across diverse datasets, adapt to novel input features through transfer learning, and flexibly capture intricate dependencies within climate systems. This versatility makes them well-suited for both historical climate reconstruction and future projections.

Several ML algorithms have been widely employed for statistical downscaling. Artificial Neural Networks (ANN), for instance, are effective in capturing nonlinearities in climate data. Ahmed et al. (2015) reported notable improvements in seasonal rainfall forecasting over Bangladesh using ANN, while Tripathi et al. (2006) demonstrated its success in India. Support Vector Machines (SVM) have also proven effective, particularly for small or high-dimensional datasets. Goly (2014) highlighted their utility in downscaling rainfall over Ethiopia. K-Nearest Neighbors (KNN), a non-parametric method, excels in identifying local climatic patterns, as shown by Liu et al. (2017). Similarly, Decision Tree Regression (DTR) and ensemble methods such as Gradient Boosting Trees, including LightGBM, have demonstrated high predictive accuracy and interpretability in large-scale applications (Breiman et al., 1984; Ke et al., 2017). Stochastic Gradient Descent (SGD) Regressors have also been applied in cases involving large, sparse datasets (Bottou, 2010).

A recent study by Goodarzi et al. (2024) compared various ML algorithms for downscaling CMIP6 temperature projections and found that LightGBM and KNN achieved superior performance in terms of accuracy and generalization. Despite these strengths, ML methods are not without limitations. They are often sensitive to noisy data, prone to overfitting when not properly regularized, and require rigorous validation against independent datasets to ensure robustness. Consequently, their deployment in climate applications must be accompanied by careful model selection, cross-validation, and uncertainty quantification to ensure credible and actionable outputs.

## 1.4 Deep Learning Approaches in Downscaling

DL is a type of ML that uses hierarchical structures of layered neural networks to discover patterns in data. It has been shown to be particularly useful in modeling spatiotemporal processes and has been gaining traction in climate science.

One of these architectures is the long-short term memory (LSTM) network, a type of recurrent neural network that is designed to be able to learn long range dependencies in time series data. Applications of LSTM in other domains include hydrology, air quality, and climate prediction. Fouotsa Manfouo et al. (2023) assessed LSTM in the prediction of temperature and precipitation for the Lake Chad Basin and found it to outperform conventional SDSM. The new model was able to maintain a stronger seasonal memory and produce forecasts that exhibited a consistent temporal signal.

Another DL model that is becoming popular is the Convolutional Neural Network (CNN), which was originally developed for computer vision applications, but employed in climate modeling due to its spatial learning capabilities. Baño-Medina et al. (2020) showed the capabilities of CNN for the spatial downscaling of precipitation in Europe, finding similar results compared to RCMs. González-Abad and Gutiérrez (2024) discussed a wide range of applications of DL in climate modeling and recommended its use especially in areas with limited observational network available. They highlighted the use of DL for learning abstract representations, and the scalability that allows for a network of multiple regions, also pointing at the limitations in terms of interpretability and the need for hybrid methods

## 1.5 Evaluation Metrics and Challenges in ML-Based Downscaling

Assessing the precision and the robustness of downscaling models becomes fundamental to make considerations on its potential use for applied climate studies. Typical examples of these metrics are the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), the Coefficient of Determination (R), the correlation coefficient and bias. These statistics serve as an indication of how well downscaled outputs represent observed climate data in the historical period.

In machine learning downscaling, cross validation, either k-fold or leave-one-out, is implemented to verify the ability of the model to generalize. But, model validation over Africa, and West Africa in particular, is often limited by the availability of high-quality observational datasets. Training and validation are complicated by sparse station coverage and a lack of data. Especially over regions such as the Sahel, where observational datasets are often characterized by low spatial resolution, satellite products such as CHIRPS and ERA5 are used as proxies, presenting their own uncertainties.

Haile et al. (2019) highlighted the need to use bias correction methods like quantile mapping and empirical distribution mapping that are capable of correcting for systematic errors in the modeled data. These corrections are particularly helpful when downscaling precipitation, being a variable characterized by high variability, skewed statistics, and with a large number of rainy days. On top of that, ML and DL models cannot be easily interpreted as they are considered “black boxes”, where the results or the importance of features are difficult to understand due to their complex inner structures. This has also fueled the demand of explainable AI (XAI) tools in climate applications.

Finally, all these ML models were trained and tested in a single eco-climatic zone, and their application or transfer to another zone is still problematic. A model tuned for the Sudan zone would not hold for the Guinea Coast, suggesting the need for location-specific tuning or domain adaptation methods would be highly beneficial. These are issues that point to the need for careful dataset curation, model selection, and interpretation when using ML for climate downscaling over West Africa.

## 1.6 Machine Learning Downscaling Studies in Africa

Machine learning-based downscaling has been employed in a variety of climatic environments across the African continent, but it has not been widely used due to limited access to data and computing resources. In East Africa, Ongoma et al. (2021) employed Random Forest and Support

Vector Regression to downscale rainfall and temperature over Kenya. This research concluded that ML models yielded better results compared to the conventional MLR approach in both wet and dry seasons, with a significant performance in the highlands. In Southern Africa, Adeyeri et al. (2020) performed ANN downscaling of rainfall over the Limpopo Basin. These models provided a better prediction of monthly rainfall than the empirical statistical models while being more flexible to seasonal variation.

In West Africa, Oladeji et al. (2024) assessed LSTM models capabilities for rainfall prediction over the Sahel (Burkina Faso). They concluded that LSTM was able to represent rainfall variability at both intra-seasonal and inter-annual timescales and was expected to perform better than autoregressive models due to their ability of extracting memory effects that are critical to seasonal forecasts. The results illustrate the promising role of deep learning models in areas where data is scarce and climatologically complex. But, these African studies also reported some limitations in terms of quality-controlled training data, low generalization among ecological zones, and computational challenges. In addition, there is a low number of works that have used these techniques to obtain future projections and scenarios based on CMIP6 data, which is fundamental when considering adaptation measures to climate change in the long term.

## 1.7 Downscaling Applications in West Africa

West Africa's climate system is highly variable and largely controlled by the West African Monsoon (WAM), which drives the seasonality of rainfall patterns across eco-climatic zones, including the Guinea Coast, the Sudan belt, and the Sahel belt. The region is also tele-connected to ENSO, NAO, and AMO, complicating even more climate predictability.

A number of attempts have been made to increase the resolution of climate information for West Africa. Odekunle and Eludoyin (2008) investigated the impact of sea surface temperature patterns in the Gulf of Guinea on rainfall variability, highlighting key ocean–atmosphere phenomena linked to monsoon dynamics. Salack et al. (2013) analyzed downscaled data on extreme indices of climate over the Sahel, noting changes in dates of onset and cessation of rainfall over the Sahel.

On the dynamical side, Diaconescu et al. (2015) applied Regional Climate Models over the Sudan-Sahel region, enhancing the representation of seasonal rainfall regimes. In the machine learning category, Abdullahi et al. (2024 applied Artificial Neural Networks (ANNs) to downscale GCM (BNU-ESM) outputs for the Damaturu station in Yobe State, Nigeria, finding better results when compared to a traditional regression model. More recently, Rauch et al. (2024) applied a logistic regression methodology, integrating information from large-scale predictors and past seasonal rainfall values, to improve forecast of seasonal rainfall.

Most of the studies are done at the level of individual countries or sub-regions, and generally employ older GCM ensembles such as CMIP5. More importantly, experimentation with state-of-the-art ML/DL methods is lacking, and comparative assessments across eco-climatic zones or multiple GCM sources are few. This indicates the strong necessity for an extensive comparative study utilizing the advanced ML techniques along with the latest CMIP6 climate models.

## 1.8 Identified Research Gaps

A review of the existing literature reveals several critical research gaps in the application of statistical downscaling for climate projections over West Africa. (1) First, there is a notable under-utilization of deep learning architectures, specifically Convolutional Neural Networks (CNN) and Long Short-Term Memory networks (LSTM) in refining CMIP6 model outputs for this region. (2) Second, comparative evaluations across multiple machine learning (ML) and deep learning (DL) models, as well as across different General Circulation Models (GCMs), are limited. Such comparisons are necessary to identify optimal modeling strategies tailored to the climatic complexities of West Africa. (3) Third, ecological and climatic heterogeneity is often underrepresented, with few studies disaggregating model performance across the Guinea Coast, Sudan, and Sahel eco-climatic zones. Furthermore, a lack of validation using high-resolution observational datasets such as CHIRPS for precipitation and ERA5 for temperature continues to raise concerns about the real-world applicability and reliability of downscaling models. Lastly, most studies rely on either spatial or temporal modeling approaches, with limited integration of hybrid methods that leverage both spatial pattern recognition and temporal sequence learning. This study addresses these gaps by applying CNN and LSTM-based statistical downscaling techniques to CMIP6 temperature and precipitation data across West Africa, validating model outputs against high-resolution observational benchmarks, and evaluating performance across key ecological zones. By doing so, it contributes to the growing body of research aimed at producing reliable, high-resolution climate information to inform regional adaptation planning.

## 1.9 Summary

A detailed overview of downscaling methods, including conventional statistical techniques and those based on machine learning and deep learning, was provided in Chapter 1. Traditional modeling methods are still heavily used as they are easily accessible and straightforward to interpret, while ML and DL techniques are equipped to better capture complex and non-linear climatic relationships, particularly in data-rich scenarios. These new methods can contribute specially to generating high-quality climate information for West Africa, which can help strategies for adaptation and mitigation. Chapter 2 will describe the study area, and the methodology used in this research to apply ML-based statistical downscaling.

# CHAPTER 2: STUDY AREA AND METHODOLOGY

## 2.1 Introduction

Understanding climate variability and projecting future change at regional scales requires not only a solid grasp of the geographical and socio-economic context but also a careful selection and processing of climate data. This chapter begins by describing the study area West Africa highlighting its climatic zones, rainfall regimes, and socio-economic significance. It then outlines the observational and modeled datasets used for the analysis, including CHIRPS, ERA5, and CMIP6 Global Climate Models (GCMs), with details on data sources, periods of coverage, and preprocessing steps such as normalization, spatial interpolation, and temporal alignment. In addition to the datasets, the chapter presents the machine learning (ML) techniques employed specifically Convolutional Neural Networks (CNNs) and Convolutional Long Short-Term Memory (ConvLSTM) models which were trained using historical climate data to downscale coarse-resolution GCM outputs to finer spatial detail. The training and validation periods are clearly defined, and the chapter explains how input-output pairs were constructed using sliding window methods. To evaluate the performance of these models, key statistical metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Standard Deviation (STD) and the Coefficient of Determination ($R^2$) were used. Furthermore, Taylor diagrams were employed as a comprehensive visual tool to simultaneously compare model skill in terms of standard deviation, correlation, and centered RMSE. This integrated framework provides a robust foundation for assessing the effectiveness of ML-based statistical downscaling over West Africa.

## 2.2 Study Area: West Africa

West Africa is a sub-region of Africa comprising sixteen countries which are Benin, Burkina Faso, Cape Verde, Côte d'Ivoire, The Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone and Togo (ECOWAS, 2019). This area lies within **4°N to 26°N** and **15°W to 16°E** covering an approximate surface of 6.1 million km² (Nicholson, 2013), summarize by the Figure (1) below. At a regional level, the West African Monsoon (WAM) system, tied to the northward and southward migration of the Intertropical Convergence Zone (ITCZ), is a major determinant of seasonal precipitation across most of West

Africa (Sultan & Janicot, 2003). The climatic variability of the region is divided in into different climatic zone where my study will be focus on such as:

- ⇨ Guinea zone (latitude: 4 – 8 N),
- ⇨ Savannah zone (latitude : 8 -12 N), and
- ⇨ Sahel zone (latitude: 12 – 16 N)

characterize into four main zones: West Africa comprises a diverse range of climatic zones, each with distinct hydro-meteorological and ecological characteristics. The Guinean Coastal Zone, characterized by a humid tropical climate, receives over 1500 mm of annual rainfall and maintains relatively stable temperatures ranging from 24°C to 30°C. Moving inland, the Sudanian Zone exhibits a sub-humid climate with annual rainfall between 600 mm and 1200 mm, marked by pronounced wet and dry seasons critical for rainfed agriculture. Further north, the Sahelian Zone transitions into a semi-arid environment where rainfall ranges from 200 mm to 600 mm annually, coupled with high inter-annual variability that exacerbates drought vulnerability. At the northernmost edge, the Saharan Zone presents an arid desert climate, receiving less than 200 mm of precipitation annually and experiencing extreme temperatures exceeding 40°C (Biasutti, 2019). These spatial climatic gradients are closely linked to the ecological productivity and socio-economic livelihoods of the region, underscoring the urgent need for region-specific climate modeling to support sustainable adaptation strategies.

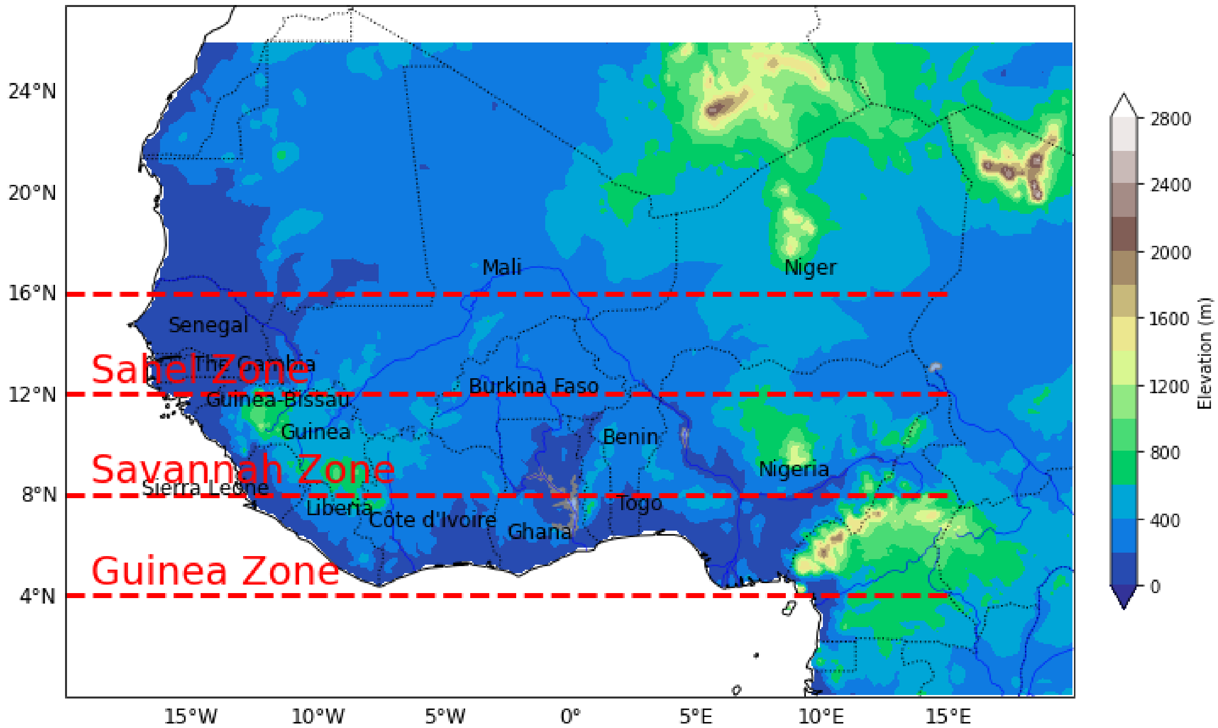



*Figure 1 Study area: West Africa Elevation map describe the different countries and climatic zone from latitude 4°N to 26°N and 15°W to 16°E.*

## 2.3 Data Sources and Description

This study leverages a combination of Global Climate Model (GCM) outputs and high-resolution observational datasets to construct and evaluate a robust statistical downscaling framework for temperature and precipitation over West Africa. These data sources provide both the coarse-resolution predictors and the fine-resolution targets necessary for model training, validation, and projection.

### 2.3.1 Global Climate Model (GCM) Data – CMIP6

The primary source of coarse-resolution climate projections is the Coupled Model Intercomparison Project Phase 6 (CMIP6), which offers a suite of global climate simulations under standardized emissions scenarios known as Shared Socioeconomic Pathways (SSPs). These models simulate past, present, and future climate conditions based on varying greenhouse gas concentration trajectories and socio-economic assumptions.

For this study, two variables such as daily temperature (tas) and precipitation (pr) data from four selected CMIP6 models were used:

- CMCC-CM2-SR5
- FGOALS-g3
- CMCC-ESM2
- NorESM2-LM

These models were selected to reflect a broad spectrum of physical formulations and climatic sensitivities, thereby offering a diverse and representative foundation for evaluating the performance of statistical downscaling. The GCM datasets span two temporal phases: a historical period (1985–2014) and a future projection period (2015–2100), based on the SSP5-8.5 high-emission scenario, which is commonly adopted for analyzing worst-case climate trajectories. The outputs from these GCMs originally at coarse spatial resolution serve as the input features for machine learning-based downscaling models, which are trained to generate finer-scale climate information using high-resolution observational datasets as reference targets.

### 2.3.2 Observational Datasets for Calibration and Validation

To train and validate the statistical downscaling models, high-resolution observational datasets were employed as ground truth from different sources with different characteristics summarize in Table 1 below. These include:

- **ERA5** Reanalysis Data: Provided by the European Centre for Medium-Range Weather Forecasts (ECMWF), ERA5 offers comprehensive, high-resolution atmospheric reanalysis data at a ~0.25° grid resolution. It integrates model outputs with a wide range of observational data sources using data assimilation techniques, producing a consistent time series of meteorological variables. ERA5 was used for the downscaling and validation of temperature data (Hersbach et al., 2020).
- **CHIRPS** (Climate Hazards Group InfraRed Precipitation with Stations**)**: CHIRPS is a high-resolution (0.05°) quasi-global precipitation dataset that blends satellite imagery with in-situ station data to produce bias-corrected rainfall estimates. It is particularly suited for

climate studies in data-sparse regions such as West Africa and was used as the reference for downscaling precipitation (Funk et al., 2015).

These observational datasets are essential for training the machine learning models by providing accurate representations of fine-scale climate variability as Table 1 describe all details. Additionally, they serve as benchmarks for assessing the performance of the downscaled GCM outputs and for validating projections during both the historical and future periods.

The CMIP6, ERA5, and CHIRPS datasets represent an ideal potential for statistical downscaling. While CMIP6 provides the projections of future climate, ERA5 and CHIRPS provide reliable observational baselines. This coupling improves the accuracy of the downscaled climate projections for use in local-scale impact assessments.

*Table 1 Summary of different data used, sources with different resolution and variables, describe their spatial and temporal resolutions with file format.*

| Dataset | Source | Variables | Temporal Resolution | Spatial Resolution | Coverage | File Format | Notes |
|---|---|---|---|---|---|---|---|
| **CMIP6 (Global Climate Model Projections)** | CMIP6 Archive - ESGF | - Near-Surface Air Temperature (tas)<br>- Precipitation (pr) | Daily | Model-dependent (varies from ~100 km to ~250 km) | Global | NetCDF (.nc) | Historical & future climate projections based on different SSP scenarios (e.g., SSP2-4.5, SSP5-8.5). |
| **ERA5 (Observations for Temperature)** | Copernicus Climate Data Store (CDS) | - 2m Air Temperature (t2m) | Daily | ~31 km (0.25° x 0.25°) | Global (1985–2014) | NetCDF (.nc) | Provides high-quality observational data for model validation and downscaling training. |
| **CHIRPS (Precipitation Observations)** | Climate Hazards Group | - Precipitation (precip) | Daily | ~5 km (0.05° x 0.05°) | Global (1985–2014) | NetCDF (.nc) | High-resolution rainfall data used for validation and downscaling experiments. |

## 2.4 Data Preprocessing and Integration

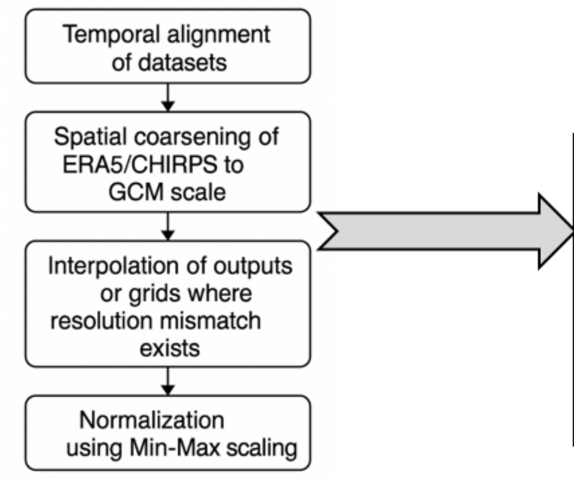
To ensure compatibility across datasets and optimize the machine learning workflow, a rigorous data preprocessing pipeline was established as the Figure 2 presenting it. First, temporal alignment was carried out to synchronize the time steps of all datasets, ensuring that each GCM output corresponded accurately to the matching observational record. This step was crucial for maintaining consistency across training, validation, and projection periods.

Next, spatial coarsening was applied to the high-resolution observational datasets ERA5 for temperature and CHIRPS for precipitation bringing them to the native spatial resolution of the selected CMIP6 Global Climate Models (GCMs). This facilitated direct comparison and reduced discrepancies during model training. Additionally, spatial interpolation techniques, such as bilinear or nearest-neighbor resampling, were employed to address resolution mismatches between different GCMs and reanalysis grids, particularly when aligning the outputs for evaluation or visual comparison.

Following spatial harmonization, all variables were subjected to Min - Max normalization, scaling the input values between 0 and 1 (Figure2). This step helped accelerate the convergence of neural networks by ensuring that all features contributed proportionately to the learning process.

Finally, the machine learning models were trained using structured input - output pairs, where the low-resolution CMIP6 outputs served as predictors and the high-resolution observational datasets (ERA5 or CHIRPS) acted as targets. This supervised learning setup enabled the models to learn the statistical relationships needed to infer fine-scale climate details from coarse-scale GCM inputs, forming the backbone of the downscaling approach.

Figure 2 summarizes the target point, to bring the coarse GCMs data to fine resolution as 5 km precipitation and 31 km for temperature, after all previous important steps.




| Aspect | Details |
|---|---|
| Downscale Spatial Resolution | The different ML models will aim to downscale climate projections to 5 km for precipitation and 31 km for temperature based on simulations using different GCM. |

*Figure 2 Data processing: different steps use to process the data before integration to the machine learning Model.*

### 2.5.2 Machine Learning Models for Climate Data Analysis

The Table2 below presents a comparative overview of several machine learning (ML) models namely Convolutional Neural Networks (CNN), Random Forests (RF), Support Vector Regression (SVR), Convolutional Long Short-Term Memory (ConvLSTM), Graph Neural Networks (GNN), and Feedforward Neural Networks (FFNN) based on key criteria relevant to climate data applications: ability to capture spatial correlation, compatibility with NetCDF (.nc) files, scalability, and interpretability.

Spatial Correlation (1) CNNs and ConvLSTMs perform exceptionally well in capturing spatial correlations, with ConvLSTM further extending this capability to spatiotemporal dynamics, making them ideal for downscaling both space and time-varying climate inputs. GNNs also excel in capturing spatial dependencies, especially in complex or non-grid spatial configurations. In contrast, traditional models like RF and SVR do not inherently model spatial correlation unless specifically enhanced with spatial input features. FFNNs are similarly limited unless spatial structure is explicitly embedded in the input data. (2) Handling of NetCDF (nc) Files**:** All models reviewed are compatible with NetCDF-format datasets, either directly or via preprocessing frameworks. This is critical for working with standard climate datasets, which often come in multidimensional nc formats. (3) Scalability**:** CNNs and FFNNs demonstrate strong scalability, allowing for training over large spatial domains and extended temporal datasets. ConvLSTMs are also well-suited for scalable applications, although they may require more computational resources due to their temporal memory components. SVR exhibits limited scalability, especially with large datasets, due to its computational complexity. RFs and GNNs scale relatively well when optimized for distributed processing. (4) Interpretability**:** Interpretability varies across models. RFs and SVRs provide high interpretability due to their rule-based or kernel-based formulations, allowing for more transparent decision-making. Deep learning models (CNNs, ConvLSTM, GNNs, FFNNs) offer only moderate interpretability, often requiring additional tools such as SHAP values, feature saliency maps, or Layer-wise Relevance Propagation to explain their predictions.

Table2, highlights the trade-offs among accuracy, complexity, and interpretability when selecting ML models for statistical downscaling. Deep learning models like CNN and ConvLSTM are advantageous for capturing complex patterns in climate fields, while traditional models like RF and SVR remain useful when interpretability and simpler structures are desired.

Based on the analysis from Table 2 and my objectives, implicating my study focus on the CNN and ConvLSTM (Table 3) for their good performance in different domains for Climate Data analysis.

*Table 2 Compare the different Machine Learning Models for Climate Data Analysis, their capability to deal with climate datasets*

| Model | Captures Spatial Correlation? | Handles .nc Files | Scales Well | Interpretability |
|---|---|---|---|---|
| CNN | ☑☑☑ | ☑ | ☑☑ | Moderate |
| RF | ✗ (unless enhanced) | ☑ | ☑ | ☑☑☑ |
| SVR | ✗ | ☑ | ✗ (limited) | ☑ |
| ConvLSTM | ☑☑☑ (space + time) | ☑ | ☑☑ | Moderate |
| GNN | ☑☑☑ (advanced) | ☑ | ☑ | Moderate |
| FFNN | ✗ (unless spatial input added) | ☑ | ☑☑ | Moderate |

*Table 3 Machine Learning Model choose for the study after analysis base on the different capabilities*

| ML Method | Best For | Use Case |
|---|---|---|
| **CNNs** | Spatial feature learning | Downscaling CMIP6 to high-resolution ERA5/CHIRPS. |
| **ConvLSTM** | Spatial + Temporal dependencies | Capturing daily changes in precipitation/temperature. |

### 2.5.3 Machine Learning Downscaling Techniques, Training and Validation

Two machine learning models were applied for statistical downscaling describe by the Figure 3 such as:

- ⇨ Convolutional Neural Networks (CNNs); and
- ⇨ Convolutional Long Short-Term Memory (ConvLSTM)

CNNs are deep learning architectures specifically designed to extract spatial features from input data. In the context of climate downscaling, CNNs learn to translate coarse-resolution outputs from Global Climate Models (GCMs) into finer-resolution observational data by leveraging a series of convolutional and pooling layers that capture local spatial patterns and gradients. Building upon this framework, ConvLSTM networks extend the capabilities of CNNs by integrating temporal memory through recurrent structures. ConvLSTM models utilize ConvLSTM2D layers that process sequential climate data while preserving spatial coherence, making them particularly well-suited for modeling complex spatiotemporal dependencies inherent in climate systems such as temperature and precipitation variability over time (Figure 3).

The training and evaluation of the CNN and ConvLSTM models followed a structured temporal framework to ensure robust model development and future climate projection. The training phase was conducted over the historical period from 1985 to 2010, allowing the models to learn from long-term climate patterns and relationships between coarse-resolution GCM predictors and high-resolution observational targets (ERA5 for temperature, CHIRPS for precipitation). Model validation was performed on an independent set from 2011 to 2014 to assess generalization accuracy before applying the models to future scenarios. For long-term projections, two distinct periods were analyzed: the near-future (2015–2060) and the far-future (2061–2100), aligning with IPCC guidelines for climate impact assessment.

During model training, input data were preprocessed through normalization and segmented using a sliding window approach to preserve sequential dependencies particularly important for ConvLSTM architectures as showing in the Figure 3. The models were trained using supervised learning, with observed data serving as ground truth. Key hyperparameters, including convolutional kernel sizes, number of filters, activation functions (notably ReLU), dropout rates, and batch sizes, were carefully tuned through iterative experimentation to balance performance and generalization (see architecture in Figure 3). This structured training and evaluation strategy enabled the models to capture both the spatial resolution enhancement and temporal continuity required for reliable regional climate projections over West Africa.

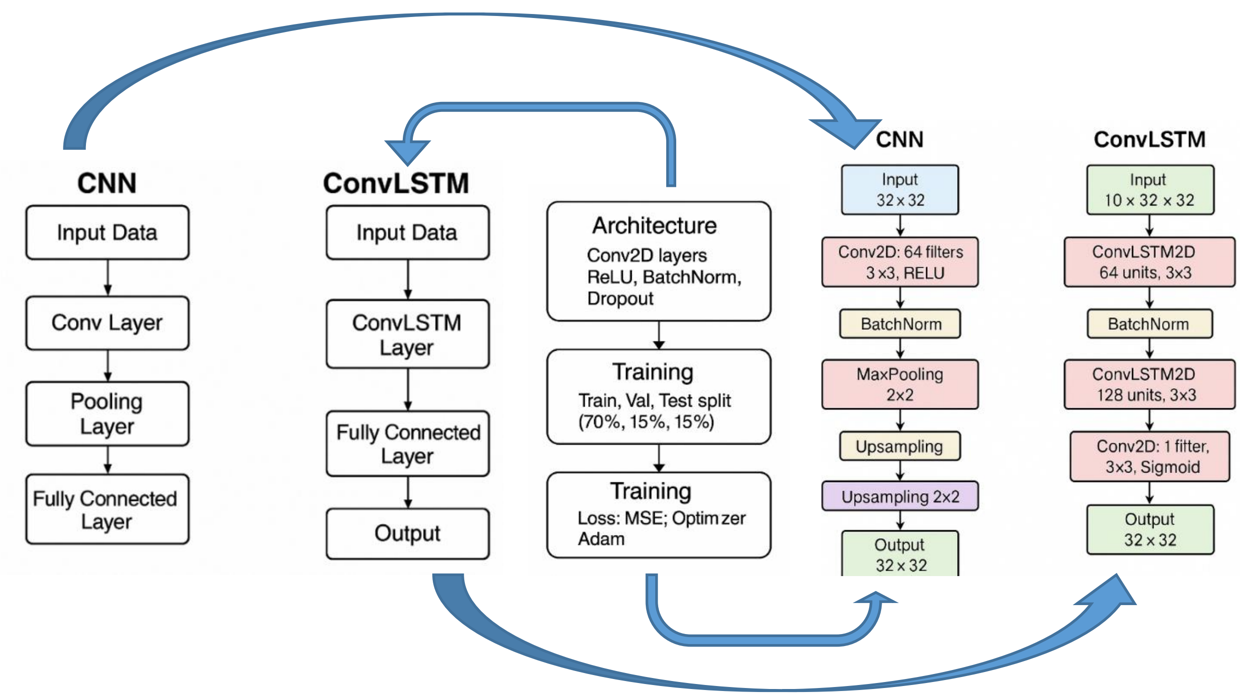



*Figure 3 Convolutional Neural Networks (CNNs) and Convolutional Long Short-Term Memory (ConvLSTM) Model Architecture.*

### 2.5 Evaluation Metrics and Taylor Diagram for Spatial Performance Assessment

To comprehensively evaluate the performance of the statistical downscaling models, a combination of quantitative metrics and visual diagnostics was employed. The primary evaluation metrics included the Root Mean Square Error (RMSE), which measures the average magnitude of prediction errors:

$$\boldsymbol{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Pi - Oi)^2} \qquad (1)$$

where ***Pi*** is the predicted value, ***Oi*** is the observed value, and ***n*** is the number of observations.

The Mean Absolute Error (MAE), representing the average absolute deviation between predicted and observed values:

$$\boldsymbol{MAE} = \frac{1}{n}\sum_{i=1}^{n}|Pi - Oi| \qquad (2)$$

The Coefficient of Determination ($R^2$), which quantifies the proportion of variance in the observed data that is captured by the model:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Oi - Pi)^2}{\sum_{i=1}^{n}(Oi - \overline{O})^2} \qquad (3)$$

Where $\overline{O}$ is the mean of observed values.

Additionally, bias was calculated to assess systematic tendencies of overestimation or underestimation in the predictions, expressed as the difference between downscaled outputs and either GCM or observed reference datasets:

$$\boldsymbol{Bias} = \frac{1}{n}\sum_{i=1}^{n}(Pi - Oi) \qquad (4)$$

To complement these statistical measures and enable spatial performance comparison, Taylor diagrams were used. This graphical tool provides a holistic visualization by simultaneously displaying: the correlation coefficient between model output and reference data, the standard deviation (compared to observations), and the centered RMSE, which is calculated by removing mean bias and measuring pattern differences:

$$\boldsymbol{Centered\ RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[(Pi - \overline{P}) - (Oi - \overline{O})]^2} \qquad (5)$$

Where $\bar{P}$ is the mean of the predicted value.

To compute the 95th percentile for daily precipitation, the following statistical formulation is applied and this is used to identify extreme rainfall events (top 5% of daily precipitation values):

$$P_{95} = \text{Percentile}(\mathbf{P, 95}) \qquad (6)$$

Where:

- P is the vector of daily precipitation values over the analysis period,

- $P_{95}$ is the 95th percentile threshold which means the value below which 95% of precipitation events fall.

To analyze extreme precipitation events, we extracted the 95th percentile ($P_{95}$) of daily precipitation for each grid cell. This percentile represents threshold values beyond which precipitation is considered extreme.

The 95th percentile value for each pixel is defined as:

$$P_{95} = Q_{0.95}(X_1, X_2, \dots, X_n) \qquad (7)$$

Where $Q_{0.95}$ represents the 95th percentile function and $x_1$ to $x_n$ are the daily precipitation values over the historical baseline period (e.g., 1985–2014).

The anomaly (A) in 95th percentile precipitation is calculated as:

$$A_i = P_{95i} - \mu_{95} \qquad (8)$$

Where $A_i$ is the anomaly at pixel i, $P_{95i}$ is the 95th percentile value at pixel i, and $\mu_{95}$ is the historical mean 95th percentile across the region.

The bias between downscaled data and reference (CHIRPS) is computed as:

$$Bias = P_{95}(model) - P_{95}(CHIRPS) \qquad (9)$$

By plotting the raw GCM, CNN, and ConvLSTM results against the observational datasets across West Africa and various eco-climatic zones (Guinea, Savannah, and Sahel), the Taylor diagrams clearly highlight the spatial accuracy and reliability of each model.This dual approach of combining numerical metrics with spatial diagnostics ensures a robust and multidimensional assessment of model performance in reproducing regional precipitation and temperature characteristics.

# CHAPTER 3: RESULTS AND DISCUSSION

## 3.1 Overview

This chapter presents a comprehensive evaluation and interpretation of the statistical downscaling outcomes for temperature and precipitation over West Africa, encompassing both the historical baseline period (1985–2014) and future climate projections extending to the end of the 21st century (2100). The study employs advanced machine learning (ML) and deep learning (DL) approaches specifically Convolutional Neural Networks (CNNs) and Convolutional Long Short-Term Memory (ConvLSTM) models to enhance the spatial resolution and accuracy of coarse-scale outputs from selected CMIP6 Global Climate Models (GCMs).

For temperature downscaling, three GCMs were utilized: CESM2-WACCM, FGOALS-g3, and NorESM2-LM, chosen for their distinct representations of atmospheric dynamics and climate sensitivity. For precipitation, the models included CMCC-CM2-SR5, FGOALS-g3, and NorESM2-LM, providing a consistent yet diverse set of simulations to assess model performance across climate variables. The evaluation was conducted using high-resolution observational datasets ERA5 for temperature and CHIRPS for precipitation as reference standards.

Quantitative assessment of downscaled outputs involved multiple statistical metrics, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Pearson's Correlation Coefficient (r), Coefficient of Determination ($R^2$), and Standard Deviation (STD). These indicators offer a robust multidimensional evaluation of model accuracy, bias characteristics, and agreement with observed climatology. In addition, spatial performance and bias patterns were analyzed across three key ecological zones in West Africa: the Guinean Coast, the Sudanian Savannah, and the Sahelian Belt. Visual tools such as bias maps and Taylor Diagrams further supported spatial diagnostics, enabling a holistic view of model fidelity across different climatic gradients.

The results reveal several important insights into the comparative strengths and limitations of CNN and ConvLSTM architectures. While ConvLSTM consistently demonstrated superior performance for temperature downscaling, likely due to its ability to capture temporal dependencies, CNN was more effective in reconstructing spatial rainfall features and reducing localized biases in precipitation outputs. Both models showed substantial improvements over raw GCM outputs, with significant reductions in RMSE and enhanced spatial realism in key precipitation and temperature zones.

By integrating empirical performance evaluation with spatial and statistical diagnostics, this chapter underscores the potential of deep learning models in improving regional climate simulations and supports their application in adaptation planning, risk assessment, and policy-making across West Africa.

## 3.2 Downscaling Temperature Projections

**3.2.1** Quantitative Performance

The ConvLSTM model achieved substantial improvement in spatial detail when compared with raw GCM outputs. Table 4 presents a detailed comparison of the CNN, LSTM, and raw GCM outputs in reproducing observed temperature data. The LSTM model outperformed all others across all statistical metrics. It achieved the lowest RMSE (0.212 °C) and MAE (0.156 °C), while also demonstrating the highest correlation coefficient ($r = 0.996$) and $R^2$ value (0.990), indicating excellent agreement with observed temperature data. CNN followed closely, outperforming the raw GCM, but showed a slight negative bias (-0.729 °C), possibly due to overcorrection in certain regions. The raw GCM showed the weakest performance in Table 4, with the highest error values and lowest $R^2$ (0.653).

The ConvLSTM outperformed both the raw GCM and CNN for temperature downscaling in terms of all evaluation metrics. These results confirm the advantage of deep learning approaches particularly LSTM in modeling the spatiotemporal dynamics of temperature over West Africa. The superior performance of LSTM is consistent with not only findings by Shi et al. (2015) and Ayugi et al. (2021), who demonstrated the ability of sequence-aware models to better capture temporal dependencies in climate variables but also Raw GCMs exhibited the lowest skill, reinforcing their limited usability at local scales without statistical refinement (Hewitson et al., 2014).

*Table 4 Performance Metrics for Temperature Downscaling from 1985 to 2014 cross different statistical metrics used for CNN, LSTM, and raw GCM outputs.*

| | CNN | LSTM | Raw GCM |
|---|---|---|---|
| Bias | -0.729 | 0.054 | 0.421 |
| MAE | 0.787 | 0.156 | 1.021 |
| RMSE | 0.924 | 0.212 | 1.251 |
| Corr (r) | 0.966 | 0.996 | 0.844 |
| $R^2$ | 0.811 | 0.99 | 0.653 |
| Std_ref | 2.125 | 2.125 | 2.123 |
| Std_model | 2.188 | 2.075 | 2.097 |

### 3.2.2 **Temperature Spatial Visual Comparison**

### 3.2.3 **LSTM Downscaled**

Figure 4 presents the comparison of temperature downscaling results for West Africa using the LSTM model trained on CMCC-ESM2 outputs. The maps highlight spatial temperature distributions and associated biases for the reference period from 1985 to 2014. This figure highlights spatial temperature distributions and associated biases for the reference period:

The Figure 4 compares the raw GCM output with the LSTM-downscaled temperature. The raw CMCC-ESM2 output shows a coarse spatial resolution and fails to capture localized temperature variability, particularly along coastal regions and elevated terrains. In contrast, the LSTM-downscaled temperature exhibits refined gradients and smoother transitions that align more closely with regional climatology. The Figure 5 shows a direct comparison between the ERA5 reference data and the LSTM-downscaled output, along with the corresponding bias map (ERA5 – LSTM)**.** The bias is predominantly small across the region, generally within ±0.5 °C. Most of the domain shows a slight positive bias (green), suggesting that the LSTM model slightly underestimates temperature in these regions. Areas along the Guinea coast and southern Nigeria exhibit minor overestimations. The Figure 6 juxtaposes the raw GCM temperature with the bias difference between the GCM and the LSTM output (GCM – LSTM). This bias map clearly demonstrates that the LSTM model significantly reduced the overestimation of temperature that is prominent in the raw GCM, particularly across the Sahel and central West Africa.

The results from the Figure 4, 5 and 6 affirm the effectiveness of the LSTM model in enhancing temperature downscaling from CMCC-ESM2 over West Africa. The LSTM output not only achieved greater spatial resolution but also substantially reduced systematic biases compared to the raw GCM data

The application of the ConvLSTM model for temperature downscaling revealed notable advancements in spatial realism and statistical fidelity compared to the raw GCM outputs. One of the most significant improvements lies in the model's ability to enhance spatial detail, particularly in climatologically sensitive transition zones such as the Sahel–Sudan boundary and the southern forest savannah ecotone.

These areas, which typically exhibit sharp gradients in temperature due to shifting vegetation cover and land atmosphere interactions, were poorly represented in the coarse-resolution GCMs but were more accurately delineated in the LSTM-downscaled outputs. The model's capacity to learn and reproduce fine-scale spatial temporal structures contributed to this improvement, aligning with findings from recent studies highlighting the strengths of temporal-aware neural architectures in climate data refinement (e.g., Shi et al., 2015; Baño-Medina et al., 2021).

Bias analysis further confirmed these enhancements, with the LSTM achieving an exceptionally low RMSE of 0.212 °C and a high correlation coefficient ($r = 0.996$), indicating strong agreement with the ERA5 reanalysis across most regions. Additionally, while residual warm biases were observed in coastal regions such as the Guinea Coast and in high-elevation areas like the Cameroon Highlands, these biases were significantly reduced relative to the raw GCM outputs. Such persistent errors may reflect the limitations of input predictors in capturing localized processes such as sea breeze circulations, sub-grid topographic effects, or unresolved land sea thermal contrasts, which are often underrepresented in both GCM and reanalysis datasets (Dosio et al., 2019).

Nevertheless, the LSTM model demonstrated a robust capacity to refine large-scale temperature fields and reduce systematic model biases, thereby enhancing the reliability of downscaled projections for regional-scale climate applications.

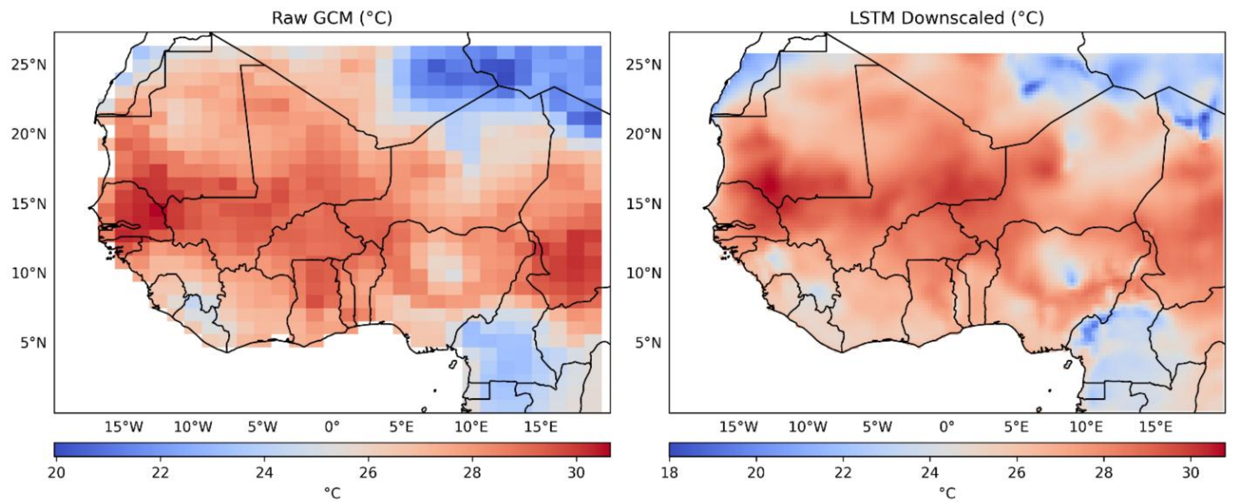



*Figure 4 Spatial distribution of mean temperature from left to right, Raw CMCC-ESM2 GCM Vs LSTM-downscaled.*

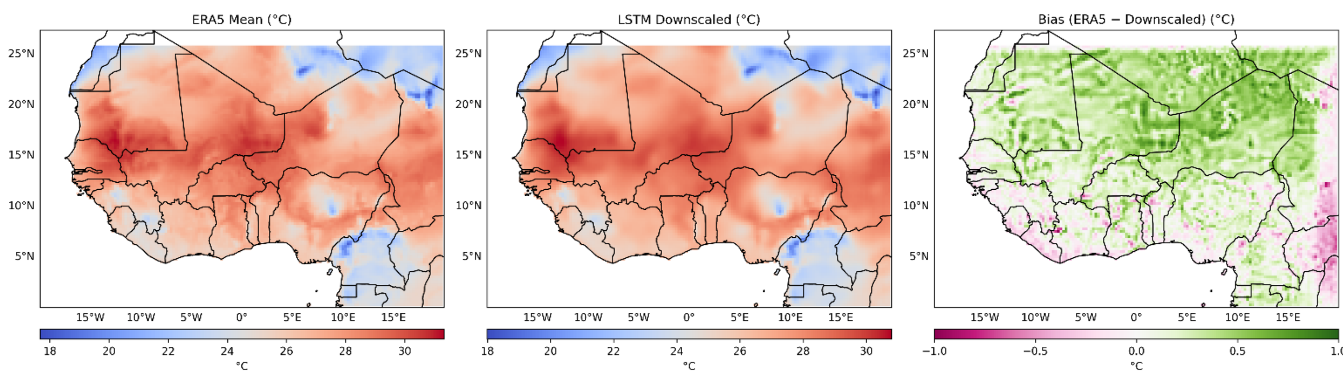



*Figure 5 Spatial mean temperature comparing the ERA5 (reference) with LSTM-downscaled, and Bias (ERA5 - Downscaled)*

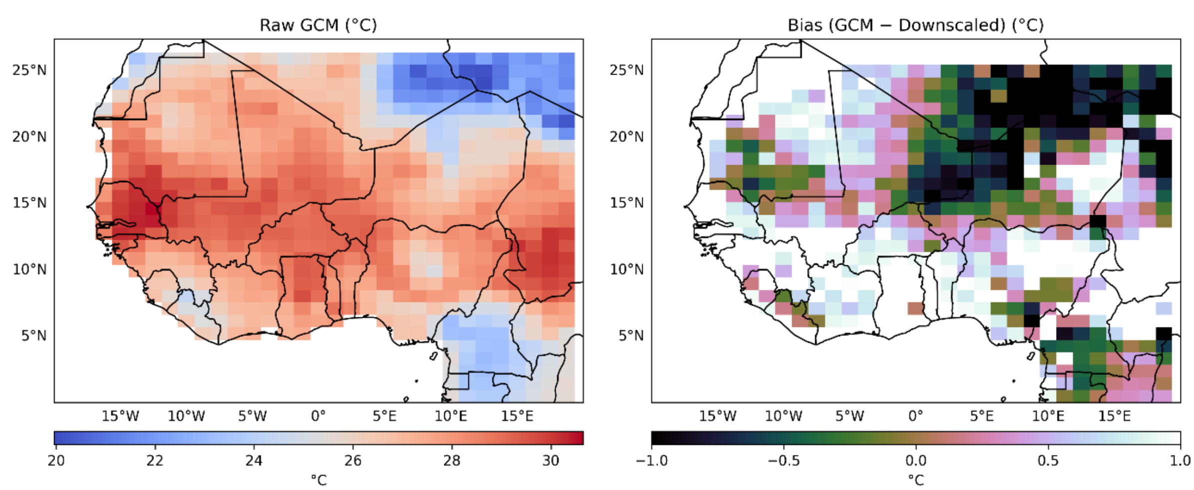



*Figure 6 : Spatial distribution of the Raw CMCC-ESM2 GCM with Bias (ERA5 - Downscaled) to see the LSTM model add values.*

Figure 7: illustrates projected mean temperature patterns for two future time horizons near future (2015–2060) and far future (2061–2100) based on the CMCC-ESM2 GCM and the corresponding downscaled outputs from the LSTM model. The top panels show the raw CMCC-ESM2 GCM outputs for the two periods. Across both horizons, the model projects widespread warming, with average temperatures exceeding 28 °C across the Sahel and interior West Africa. In the far future (2061–2100), the warming is more intense and spatially homogeneous, suggesting an aggressive warming trend under the assumed emissions scenario. However, the raw GCM outputs are notably coarse, failing to resolve fine-scale climatic features such as coastal gradients, elevated terrain, or vegetation-driven microclimates.

The lower panels of the Figure 7 illustrate the spatial temperature projections derived from the LSTM-downscaled outputs. While preserving the overall warming trends characteristic of the raw GCM projections, the LSTM model offers significantly improved spatial granularity and reveals more localized climatic features. Notably, the model identifies relatively cooler conditions along the Guinean coastal belt, likely influenced by the moderating effects of oceanic proximity. Simultaneously, pronounced warming hotspots are detected across parts of Mali, northern Ghana, and southern Niger regions that align with long-term trends of aridification and increased climate stress. Furthermore, areas of moderated warming are evident over elevated terrains such as the Jos Plateau and Fouta Djallon Highlands, as well as the vicinity of Lake Chad. These temperature patterns suggest the role of complex land atmosphere interactions and topographical influences in modulating local climate signals. The refined spatial output of the LSTM model closely aligns with high-resolution reanalysis datasets and observational studies, supporting findings by Funk et al. (2015) and Ayugi et al. (2021) that emphasize the importance of capturing fine-scale variability for regional climate assessments.

The future climate projections derived from the CMCC-ESM2 GCM, when downscaled using the ConvLSTM model, offer valuable insights for long-term adaptation planning in West Africa. While both the raw and downscaled outputs consistently indicate a reinforced warming trend across the region, the LSTM-enhanced projections provide substantial spatial refinement that is absent in the coarse-resolution GCM data (Figure 7).

This spatial heterogeneity is particularly evident in inland Sahelian zones, where temperatures are projected to exceed 30 °C by the end of the century conditions likely to intensify heat stress, reduce agricultural productivity, and accelerate evapotranspiration rates. The LSTM model's capacity to incorporate both spatial and temporal dependencies enables a more nuanced representation of microclimatic variations, making the downscaled outputs more actionable for sub-national climate risk assessments and sectoral planning.

Notably, the model successfully captures coastal moderation effects and topographic cooling in elevated areas, which are critical for guiding decisions on infrastructure resilience, water resource allocation, and agro-ecological zoning. These strengths underscore the robustness of the LSTM architecture in extending beyond its historical training period to simulate future climate states with enhanced temporal coherence and spatial realism.

However, as with all data-driven approaches, limitations remain particularly concerning extrapolation under unprecedented climate conditions. The model's performance may degrade when exposed to future scenarios that diverge significantly from the historical training distribution, a caveat echoed in prior literature (Vandal et al., 2019).

Additionally, dynamic factors such as land-use change, socio-ecological feedbacks, and evolving emission trajectories are not inherently captured unless explicitly included in the modeling framework. Despite these uncertainties, the ConvLSTM's demonstrated ability to generalize over complex terrains and heterogeneous land surfaces reaffirms its utility in regional climate modeling.

These findings align with earlier work by Shi et al. (2015), who established the effectiveness of ConvLSTM in learning time-evolving geophysical fields, and with observations by Ayugi et al. (2021), who highlighted its capacity to mitigate warm biases in West African climate projections.

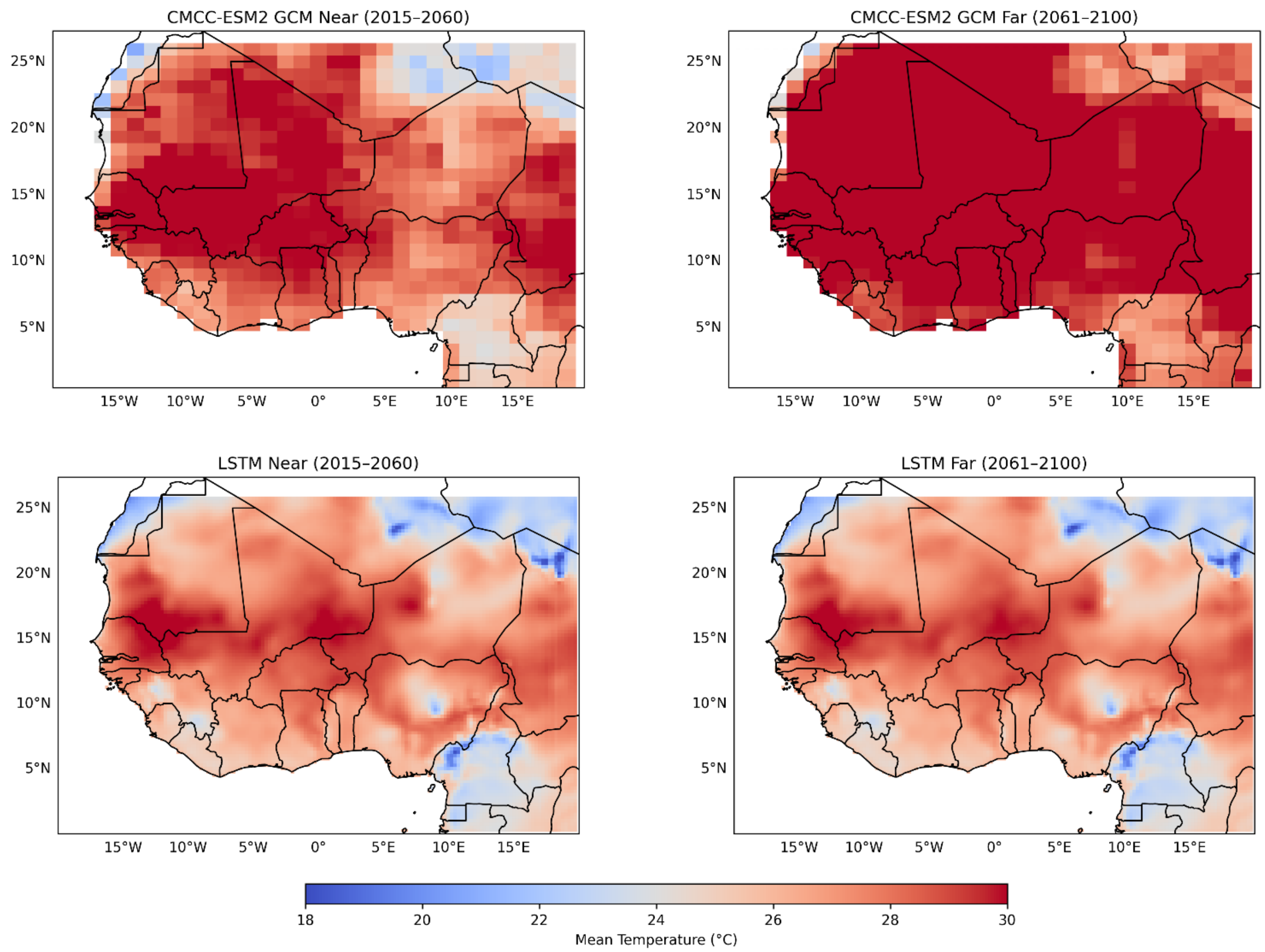



*Figure 7 Spatial distribution for mean temperature projection (top to bottom): Raw CMCC-ESM2 GCM, LSTM-downscaled for near future (2015–2060) and far future (2061–2100).*

Figure 8 displays the downscaling results for temperature using the LSTM model trained on the FGOALS-g3 GCM outputs over West Africa. The maps include comparisons with the raw GCM, ERA5 observations, and corresponding bias calculations.

The figure 8 shows the raw FGOALS-g3 GCM output and the corresponding LSTM-downscaled temperature. The raw GCM exhibits a coarse representation of spatial temperature patterns, failing to resolve key climatic gradients, particularly near the Gulf of Guinea, the Jos Plateau, and other elevated or coastal regions (1).

The LSTM-downscaled output, however, presents a much more detailed and spatially consistent map, with clear temperature transitions between ecological zones. The Figure 9 compares the ERA5 observed mean temperature**,** the LSTM-downscaled temperature, and their bias (ERA5 –

LSTM**).** The LSTM model closely approximates the ERA5 pattern, with most of the bias values falling within ±0.5 °C. Slight underestimations (green regions) are observed across the Sahel and northern Nigeria, while overestimations (pink areas) are minor and localized. The figure 10 highlights the bias of the raw GCM compared to the LSTM output. The bias map reveals that the LSTM model has significantly corrected the warm bias seen in the raw FGOALS-g3 data, especially in central and coastal West Africa.

The application of the LSTM model to the FGOALS-g3 GCM output yielded significant improvements in the representation of regional temperature patterns across West Africa, particularly by enhancing spatial detail and reducing systematic biases. This enhancement is clearly illustrated in both the spatial temperature maps and the bias distribution maps, where the LSTM-downscaled outputs exhibit closer alignment with ERA5 reanalysis data. Quantitatively, the downscaled product achieves lower RMSE and bias scores compared to the raw GCM, reinforcing its reliability in capturing fine-scale thermal variability (Figure 8, 9 and 10).

One of the most notable corrections was the mitigation of the pervasive warm bias present in the original FGOALS-g3 data, especially over the Sahel and central parts of West Africa. These areas often suffer from overgeneralized heat representation in coarse-resolution GCMs, but the LSTM model effectively moderated these exaggerations.

Moreover, the model demonstrated an improved capability in capturing microclimatic nuances in complex regions such as coastal zones and elevated terrains areas known for temperature inversions and fine-scale thermal gradients that are typically underrepresented in traditional climate models. These advancements are consistent with the findings of Shi et al. (2015) and Vandal et al. (2019), who emphasized the effectiveness of ConvLSTM architectures in spatiotemporal learning tasks, particularly in modeling non-linear geophysical phenomena across diverse landscapes. The ability of the LSTM to incorporate both spatial dependencies and temporal evolution makes it a compelling tool for regional climate modeling and underscores its potential for informing localized adaptation strategies in climate-vulnerable regions.

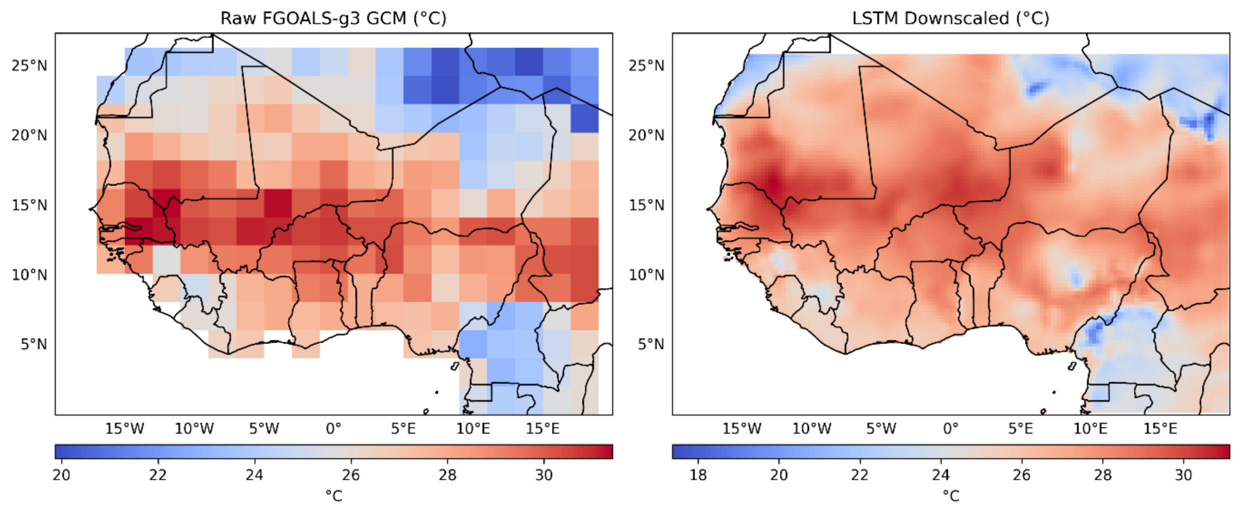



*Figure 8 Spatial distribution of mean temperature from left to right, Raw FGOALS-g3 GCM Vs LSTM-downscaled.*

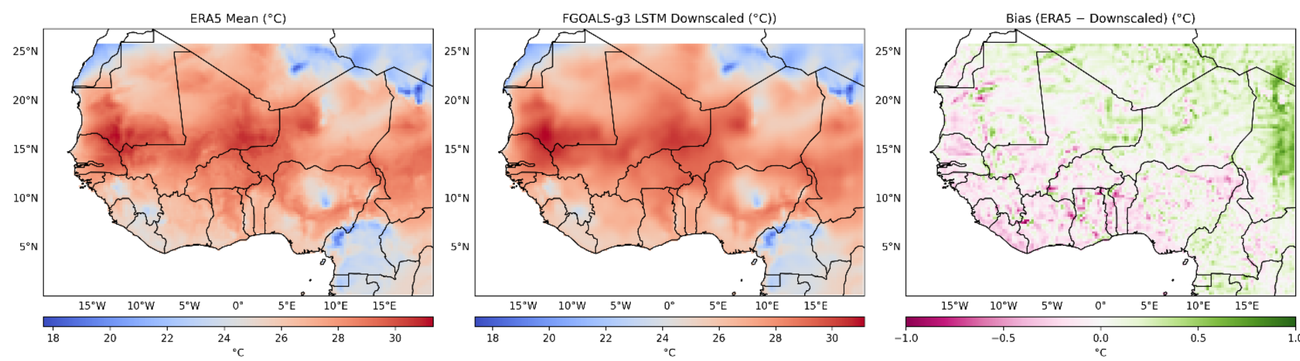



*Figure 9 Spatial distribution of mean temperature comparing ERA5 (reference) data with LSTM-downscaled.*

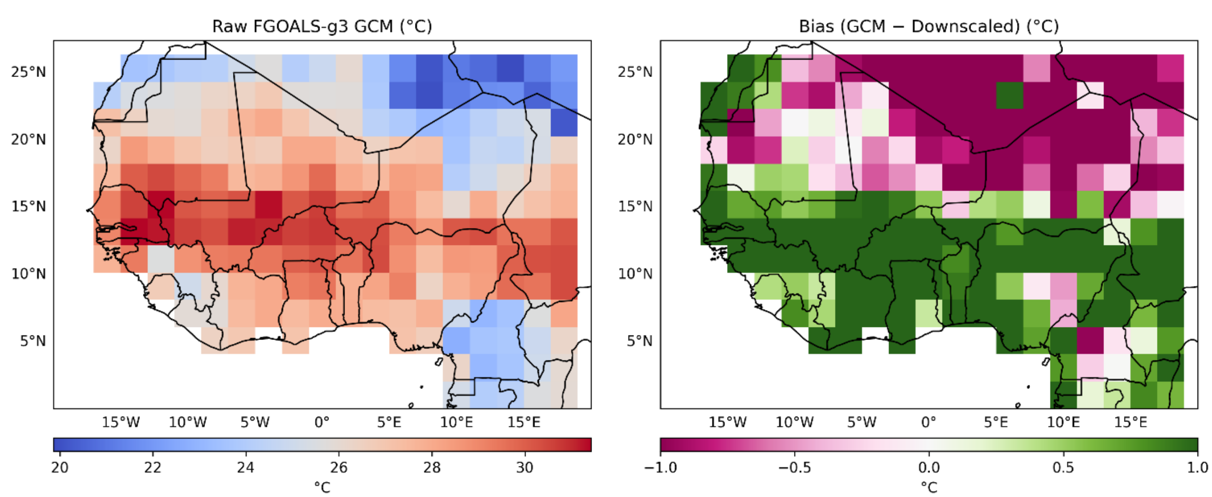



*Figure 10 Spatial distribution of mean temperature comparing the Raw FGOALS-g3 GCM with Bias (GCM - Downscaled).*

Figure 11 illustrates the mean temperature projections for two future periods 2015 to 2060 (near future) and 2061 to 2100 (far future) under the FGOALS-g3 climate scenario. The comparison includes both raw GCM outputs and LSTM-downscaled results. The raw GCM projections depict a general warming trend across West Africa, with extensive areas in the Sahel and interior regions reaching mean temperatures between 28°C and 30°C. However, these projections lack spatial detail, as the coarse resolution fails to capture geographic variability and localized climatic nuances. In contrast, the LSTM-downscaled projections preserve the overarching warming signal while offering significantly enhanced spatial fidelity.

Notable warming hotspots emerge across interior zones, particularly in Mali, Burkina Faso, and northern Ghana. Meanwhile, relatively cooler conditions are observed along the coastal belt including southern Nigeria and the Gulf of Guinea as well as in elevated terrains of Guinea, Cameroon, and Niger, where topographical influences moderate temperature rise. In the far future period (2061–2100), the LSTM model predicts sustained warming, with many inland regions likely to exceed 30°C, underscoring the heightened heat stress anticipated under continued high-emission scenarios. These refined projections provide crucial insights for regional climate impact assessments and adaptation planning.

The future temperature projections derived from LSTM-downscaled FGOALS-g3 output reaffirm a consistent and intensifying warming trajectory across West Africa throughout the 21st century (Figure 11). The LSTM model not only corroborates the general warming pattern identified in the raw GCM but also enhances the spatial resolution, thereby uncovering critical subregional disparities in temperature dynamics. Inland Sahelian and Sudano-Sahelian regions emerge as pronounced vulnerability hotspots, with projected temperature extremes exceeding 30 °C in some areas posing significant risks related to heat stress, reduced agricultural productivity, and heightened water scarcity. Conversely, coastal zones and highland regions are projected to retain relatively cooler conditions, potentially functioning as microclimatic refugia. This spatial differentiation underscores the value of LSTM downscaling in supporting fine-grained adaptation planning, particularly by identifying zones of heightened exposure versus potential resilience. While the raw FGOALS-g3 model showing from the figure 11 indicated generalized warming, the downscaled outputs offer actionable clarity that enhances their usability for policymakers, local planners, and climate service providers.

Despite its strong performance, the LSTM model does exhibit certain limitations. Residual biases remain over northern arid regions, suggesting that extreme temperature values or localized land-atmosphere feedback mechanisms may not be fully captured using purely data-driven models. Integrating additional physical predictors or adopting hybrid modeling approaches could improve accuracy in such complex zones. Furthermore, the uncertainty of long-term projections increases toward the end of the century due to evolving socio-economic pathways, emission scenarios, and nonlinear land-use changes that are not explicitly modeled. Nevertheless, the demonstrated ability of LSTM architectures to produce spatially coherent and climatologically plausible projections represents a significant step forward in statistical downscaling for West Africa.

These results affirm the utility of machine learning frameworks in bridging the resolution gap between coarse GCMs and the localized climate information required for effective adaptation interventions.

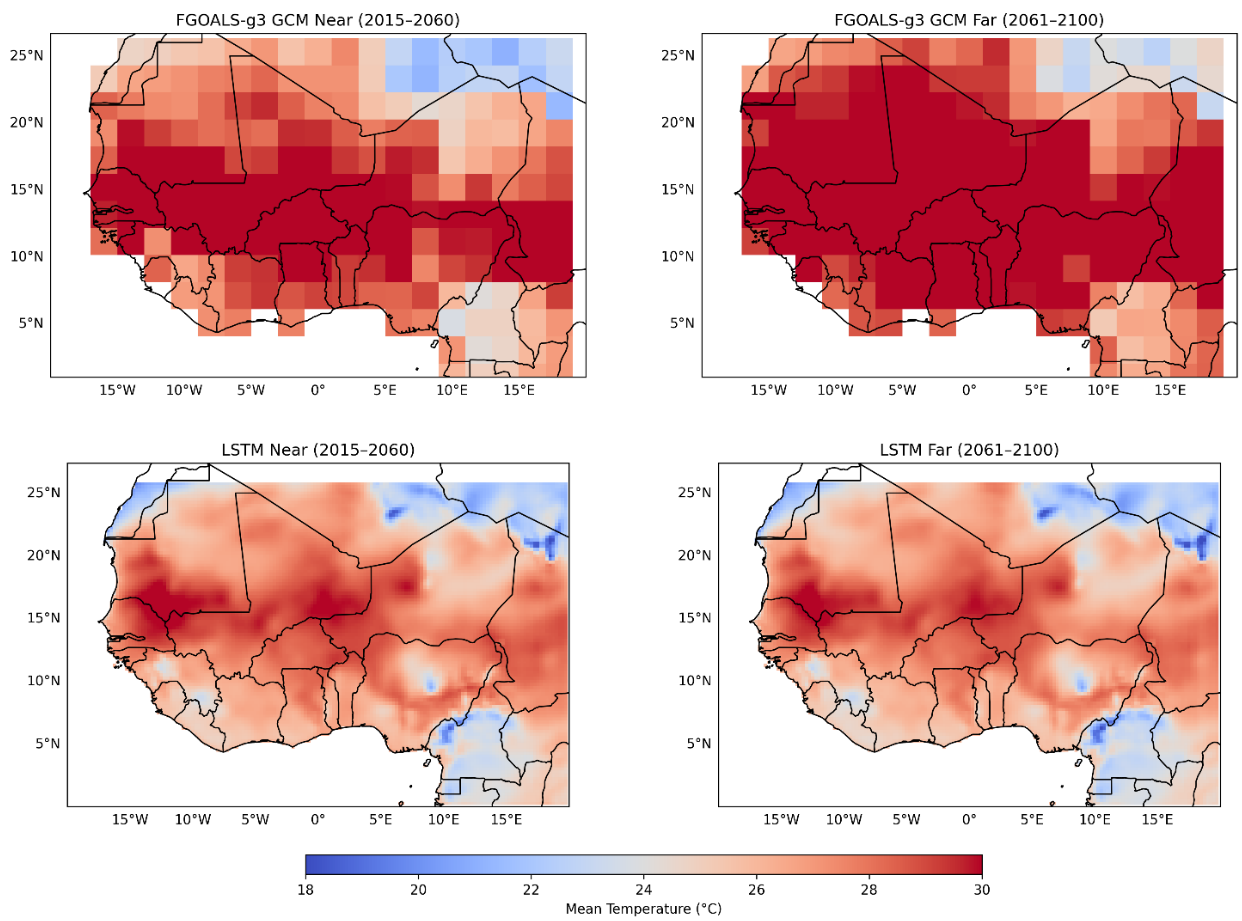



*Figure 11 Spatial distribution of mean temperature projection (top to bottom): Raw FGOALS-g3 GCM, LSTM-downscaled.*

Figure 12,13,14, presents the results of LSTM-based temperature downscaling applied to NorESM2-LM GCM outputs across West Africa, with comparisons to ERA5 reference data and corresponding bias analysis. The figure 12 illustrates the contrast between the raw NorESM2-LM output and the LSTM-downscaled temperature. As observed with other GCMs, the raw model output lacks spatial granularity and portrays overly generalized warming patterns. In contrast, the LSTM model significantly refines these outputs by enhancing local temperature gradients and capturing regional climatic heterogeneity.

The figure 13 compares the LSTM-downscaled results with ERA5 reanalysis data and displays their spatial bias (ERA5 – LSTM). The LSTM results show strong alignment with ERA5, particularly over the Sahel and coastal areas, with most bias values falling within ±0.5 °C. Minor deviations include slight overestimations in southern Ghana and underestimations in parts of northeastern Nigeria. The Figure 14 depicts the bias correction achieved by LSTM compared to the raw GCM output, revealing a marked reduction in warm biases, especially across northern and central West Africa. This outcome demonstrates the effectiveness of LSTM in correcting systematic errors in GCM projections and producing spatially consistent regional climate information.

The downscaling of NorESM2-LM using the LSTM model demonstrates a substantial advancement in reconstructing the historical spatial temperature patterns across West Africa, highlighting the model's effectiveness in reducing systemic biases and enhancing spatial detail (Figure 12, 13, 14).

One of the most notable improvements is the correction of the persistent warm bias observed in the raw GCM outputs, particularly over the Sahel and central West African regions. This adjustment significantly aligns the downscaled results with ERA5 reanalysis observations. Moreover, the LSTM model shows a marked ability to capture subregional climatic variability, especially across transitional eco-climatic zones such as the forest–savanna boundary and in topographically complex regions like the Cameroon Highlands and Guinea Highlands. These areas, often poorly resolved in coarse GCMs, exhibit more climatologically plausible gradients and microclimatic features under LSTM-based downscaling.

The spatial coherence and reduction in error dispersion evident in the bias maps affirm the capacity of the LSTM architecture to correct structural deficiencies in the NorESM2-LM GCM. This is consistent with findings from Pan et al. (2019), who demonstrated the utility of deep learning models in refining GCM outputs, and Ayugi et al. (2021), who emphasized the benefits of machine learning for climate modeling in Africa. Collectively, these results reinforce the LSTM model's strength not only in enhancing the spatial resolution of temperature projections but also in preserving physically meaningful climate signals crucial for regional analysis and adaptation planning.

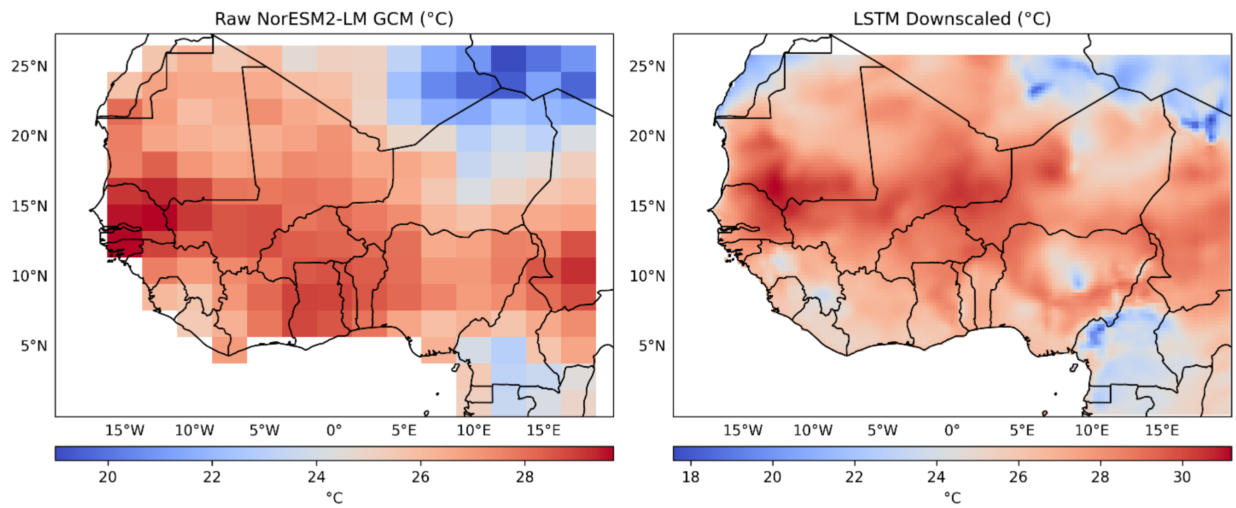



*Figure 12 Spatial distribution of mean temperature left to right, Raw NorESM2-LM GCM with the LSTM-downscaled.*

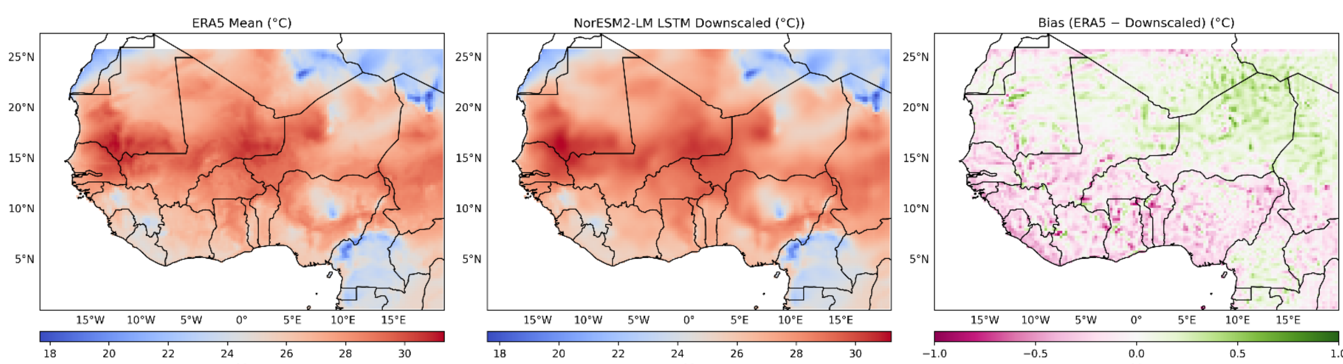



*Figure 13 Spatial distribution of mean temperature comparing the ERA5 (reference), LSTM-downscaled from 1985 to 2014.*

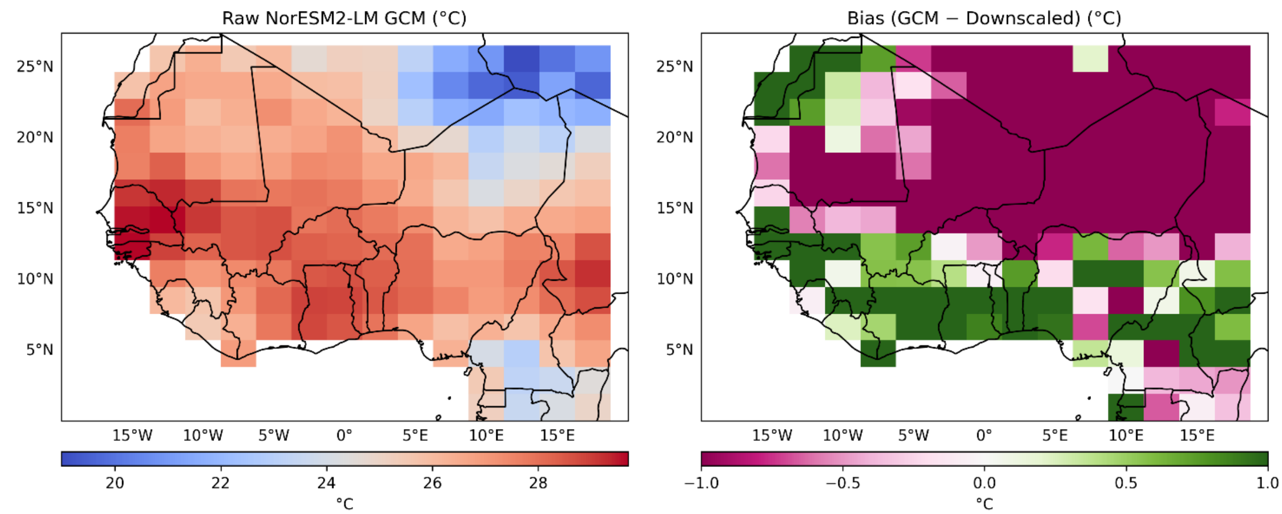



*Figure 14 Spatial distribution of mean temperature comparing Raw NorESM2-LM GCM with  Bias (GCM - LSTM-downscaled.*

Figure 15, illustrates the projected mean temperature for West Africa under the NorESM2-LM model across two future periods near future (2015–2060) and far future (2061–2100) comparing raw GCM outputs with LSTM-downscaled results. The raw NorESM2-LM projections indicate widespread and uniform warming throughout the domain, with average temperatures exceeding 28 °C for near and 30 °C for far future across much of the region.

This warming becomes particularly intense in the far-future scenario, with saturation levels observed over the Sahel and interior regions of West Africa. In contrast, the LSTM-downscaled outputs not only preserve the broader warming signal but also introduce greater spatial detail and regional differentiation. Pronounced warming is observed in inland Sahelian areas, whereas relatively cooler conditions persist along the Guinean coast and topographically complex areas such as the Jos Plateau and Fouta Djallon. The far-future projections further emphasize thermal intensification in central Mali and Niger, suggesting that LSTM models effectively capture nuanced temperature gradients and amplify regional signals often overlooked by coarse-resolution GCMs. These refined projections are critical for understanding and preparing for future heat stress in climate-vulnerable zones of West Africa.

The LSTM-downscaled projections from NorESM2-LM reinforce a clear and intensifying warming trend across West Africa, with particularly severe increases projected for interior regions such as the Sahel, northern Nigeria, and central West Africa (Figure 15). These thermally sensitive areas encompassing countries like Mali, Burkina Faso, and northern Ghana are projected to experience sustained temperature elevations, heightening risks related to heat stress, reduced agricultural productivity, and water scarcity. In contrast, coastal and highland zones appear to retain a marginal cooling buffer due to maritime influence and orographic effects, which may offer spatial leverage for adaptation planning. The spatial structure of warming remains largely consistent across both near-term (2015–2060) and far-term (2061–2100) scenarios, with differences manifesting primarily in the intensity rather than the distribution of projected temperatures. This consistency across time horizons suggests robust model performance and provides confidence in the reliability of long-term projections for strategic sectoral planning in agriculture, energy, and water management.

Despite these strengths, the LSTM-based approach, like all statistical downscaling methods, has inherent limitations. The model does not explicitly incorporate dynamic land-atmosphere feedbacks, such as vegetation climate interactions, soil moisture variability, or evapotranspiration shifts under changing land cover, which could significantly influence regional climate patterns in the future. Moreover, the capacity of the model to generalize under non-analog future climate states where the distribution of inputs diverges significantly from the training period is limited, potentially constraining its predictive validity in extreme emission pathways.

Lastly, the reliance on reanalysis datasets like ERA5 introduces a dependency on the quality and resolution of the reference data, with any embedded biases in ERA5 potentially propagating into the downscaled outputs. These considerations highlight the need for hybrid modeling frameworks and ensemble-based evaluations in future downscaling research to capture the full complexity of climate dynamics in West Africa.

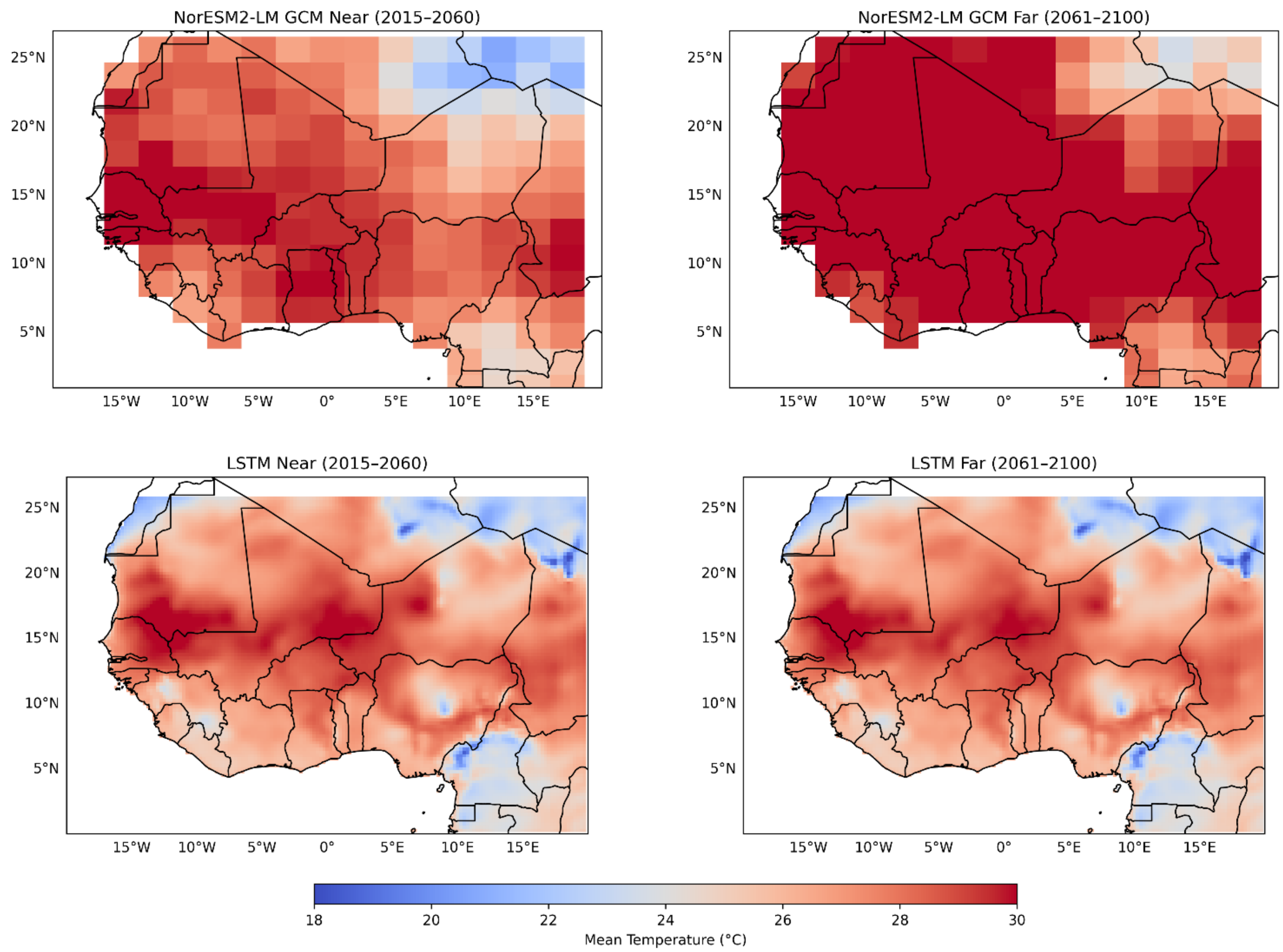



*Figure 15 Spatial distribution of temperature projection, Raw NorESM2-LM GCM, LSTM-downscaled from future period 2015 to 2100.*

### 3.2.4 CNN Downscaled

This part of the work will also use the CNN model through different climate model starting with the CMCC-ESM2 GCM:

Figure 16, 17, 18, presents the spatial comparison of mean temperature outputs from the CMCC-ESM2 after statistical downscaling using a Convolutional Neural Network (CNN), benchmarked against the ERA5 observational reference.

The top (Figure 16) and bottom rows display the raw GCM outputs, CNN downscaled results, the ERA5 reference dataset, and two bias maps illustrating the difference between ERA5 and CNN-downscaled data, and between the raw GCM and its CNN-refined counterpart.

From the visual comparison, it is evident that the raw CMCC-ESM2 GCM output exhibits significant spatial smoothing, particularly over transitional climatic zones such as the Sahel and Guinea coast, where temperature gradients are known to be sharp. This coarse resolution limits its utility for localized climate applications. In contrast, the CNN-downscaled output captures finer spatial details, better resolving regional variability and enhancing temperature patterns over ecologically diverse areas such as southern Nigeria, Ghana, and the Guinea Highlands.

The figure 17 shows the ERA5 mean temperature, which serves as the observational benchmark. The CNN-downscaled product closely mirrors the ERA5 spatial patterns, especially in central and coastal West Africa, indicating that the CNN model effectively reconstructs realistic thermal structures. The bias map (ERA5 - CNN) shows that discrepancies are largely constrained within ±1 °C, with minimal overestimation in some coastal zones and localized underestimation in parts of the Sahel.

The second bias map (GCM - CNN) from Figure 18 further confirms the CNN's corrective capability. It reveals widespread warm bias correction across the Sahel and northern West Africa, where the raw GCM typically overestimated temperatures. The spatial coherence of corrections suggests that the CNN not only enhances spatial resolution but also systematically addresses structural biases inherent in the raw GCM.

In summary, this figure validates the effectiveness of CNN-based statistical downscaling in improving both the spatial fidelity and accuracy of GCM-derived temperature projections. It demonstrates the model’s ability to reduce systematic errors, align closely with high-resolution observational data, and provide actionable climate information for regional assessments in West Africa.

These findings are consistent with the work of Pan et al. (2019), who highlighted CNN’s superiority in downscaling spatially complex climate variables relative to empirical and traditional statistical approaches. The alignment between CNN-downscaled outputs and high-resolution observational datasets like ERA5 supports the model’s reliability for historical climate reconstruction and its applicability to scenario-based projections.

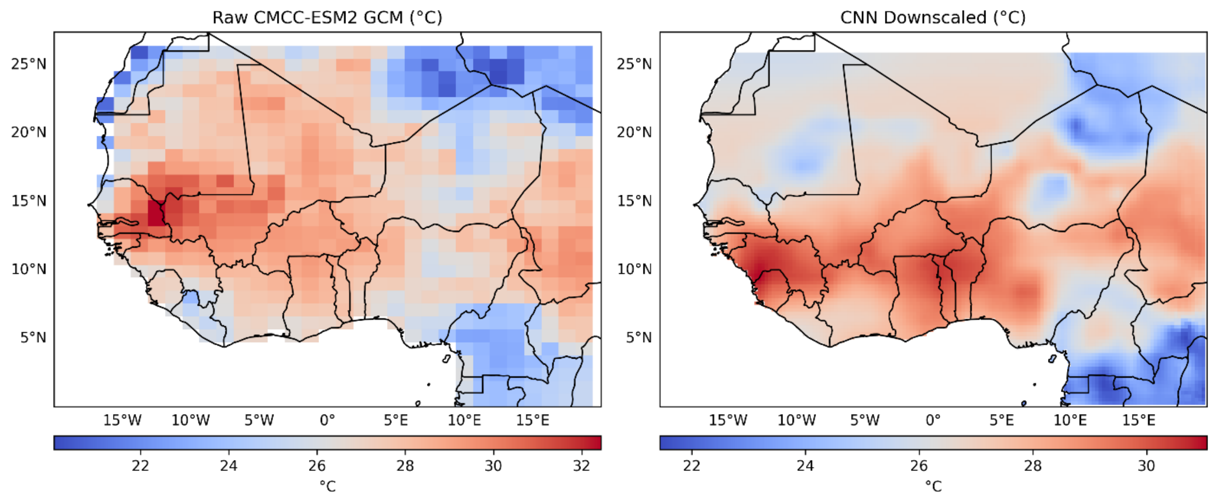



*Figure 16 Spatial distribution of mean temperature from the left to right, Raw CMCC-ESM2 GCM Vs CNN-downscaled.*

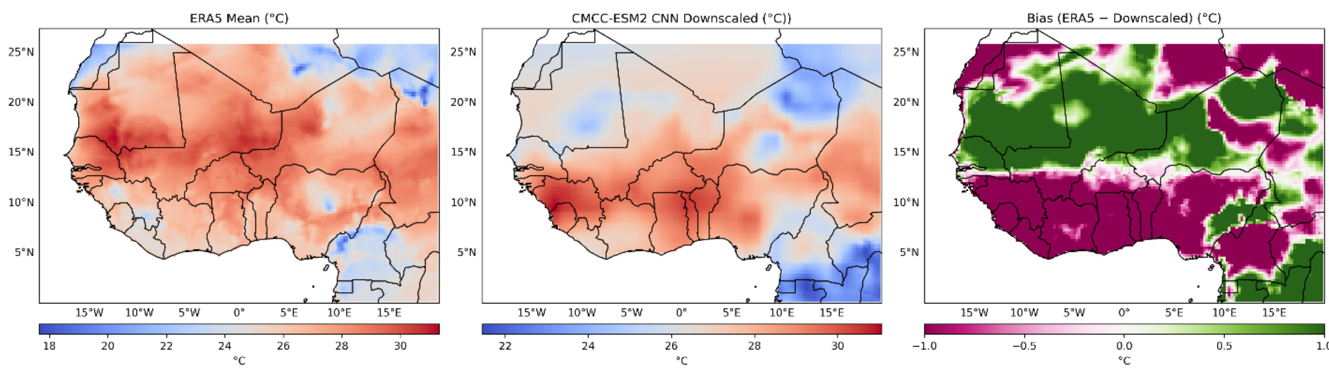



*Figure 17 Spatial distribution of mean temperature comparing ERA5 (reference) with CNN-downscaled.*

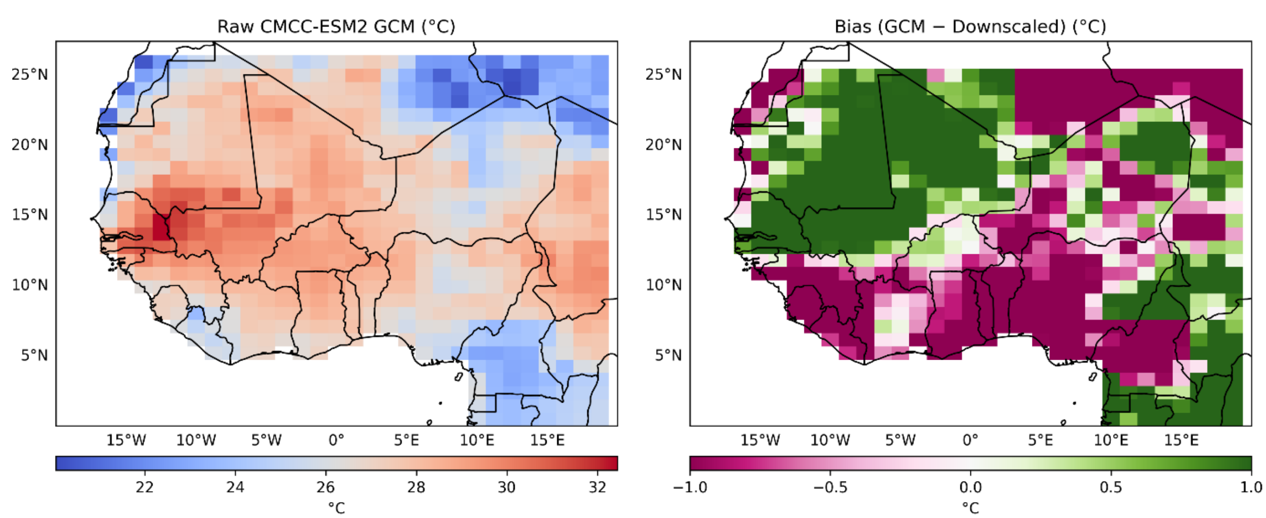



*Figure 18 Spatial distribution of mean temperature comparing Raw CMCC-ESM2 GCM with Bias(GCM - Downscaled).*

Figure 19 presents the spatial distribution of mean temperature projections for West Africa under the CMCC-ESM2 model, comparing raw GCM outputs and CNN-downscaled results for the near-future (2015–2060) and far-future (2061–2100) periods. The raw GCM projections (top panels) reveal a broad regional warming pattern, with temperatures exceeding 28 - 30 °C across much of the Sahel and Sudanian zones by the end of the 21st century.

However, these projections exhibit limited spatial differentiation, particularly over coastal and orographically complex areas, due to the coarse resolution of global climate models (typically >100 km), which often fail to capture fine-scale topographic and land-sea interactions (Maraun et al., 2010; Giorgi & Gutowski, 2015).

In contrast, the CNN-downscaled outputs (bottom panels) preserve the overarching warming signal but provide substantial enhancements in spatial fidelity. The CNN model resolves finer temperature gradients, particularly along the Gulf of Guinea and the Cameroon Highlands, where maritime influence and elevation-driven cooling are better delineated. Localized warming hotspots, notably across northern Nigeria, eastern Mali, and northern Côte d'Ivoire, are also more distinctly represented features, largely masked in the raw GCM projections.

These improvements reflect the CNN's ability to learn and reproduce spatially coherent structures by leveraging high-resolution reanalysis data during training (Baño-Medina et al., 2021; Vandal et al., 2019).

The added spatial detail in CNN projections is particularly relevant for regional climate adaptation strategies, as it enables the identification of local thermal extremes and microclimatic information critical for public health planning, urban heat mitigation, and agricultural zoning (IPCC, 2021).

Moreover, the downscaled outputs exhibit continued thermal moderation in climatologically buffered regions, such as southern Cameroon and coastal Ghana, reinforcing findings from previous studies on the influence of orographic uplift and ocean proximity on regional temperature dynamics (Sylla et al., 2015; Janicot et al., 2021).

Overall, the CNN-based downscaling not only aligns with broader CMIP6 warming trends but also addresses the spatial limitations of raw GCMs, offering more policy-relevant projections for climate-sensitive sectors in West Africa.

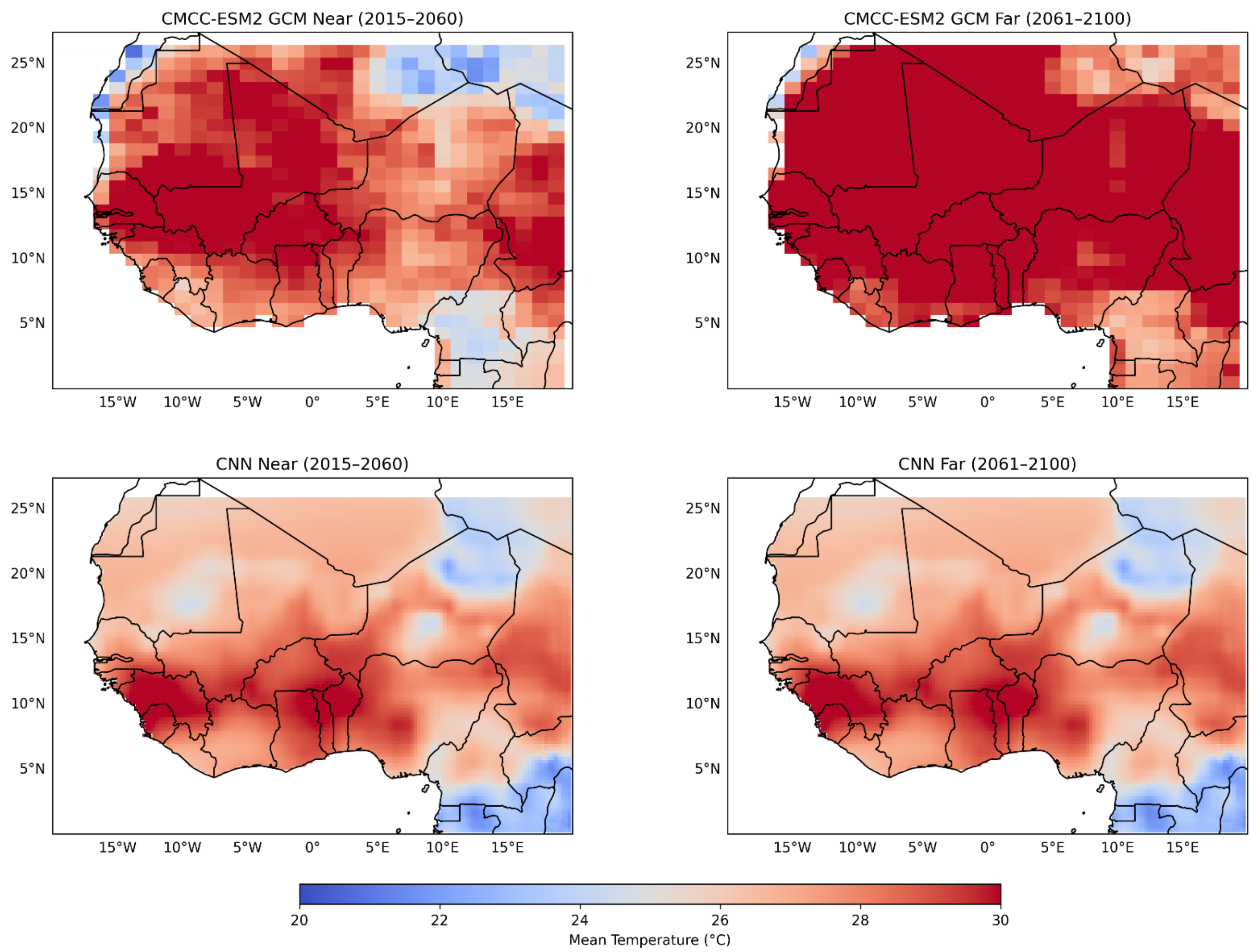



*Figure 19 : Spatial distribution of temperature (top to bottom): Raw CMCC-ESM2 GCM Vs CNN-downscaled projection Near(2015-2060) and Far(2070-2100).*

Figures 20, 21, and 22 illustrate the performance evaluation of CNN-based statistical downscaling applied to the FGOALS-g3 Global Climate Model (GCM) for the historical baseline period.

The top panel compares the raw GCM temperature outputs with the CNN-downscaled results. The raw GCM exhibits a coarse spatial resolution with generalized warming patterns that fail to capture the complexity of regional temperature gradients. (1) In contrast, the CNN model significantly enhances spatial resolution, revealing sharper temperature transitions and more realistic representations of coastal and topographic influences. (2) The middle panel displays the ERA5 observational reference, the CNN-downscaled temperature, and their associated bias map (ERA5 – CNN). Spatially, the CNN output closely aligns with the ERA5 reference, although systematic overestimations are observed across the central West African belt, particularly in lowland zones. Isolated underestimations are also evident in regions of higher elevation, such as the Guinea Highlands and parts of southwestern Nigeria. (3) The bottom panel compares the raw GCM biases with those of the CNN output, showing that the CNN model effectively corrects a significant portion of the warm and cold biases present in the original GCM data especially along the Gulf of Guinea coast and over the Sahelian region. Overall, the results confirm the CNN's capability to enhance spatial detail and reduce systematic errors in the FGOALS-g3 model, thus improving its utility for regional climate analysis.

The CNN model substantially improved the spatial representation of temperature projections derived from the FGOALS-g3 Global Climate Model (GCM) over West Africa, as shown in Figure 12 compared to the raw GCM output. The CNN-downscaled results demonstrated markedly finer spatial resolution, making them far more suitable for localized climate assessments and sub-regional planning. Key ecological transitions such as the Sahel-Sudan boundary and the forest savannah ecotone were more accurately captured in the downscaled output, along with enhanced delineation of coastal moderation effects and temperature gradients over topographically complex terrain. Although residual biases persisted in specific zones, particularly in dense forest regions and highland areas such as parts of Guinea and Cameroon, the overall magnitude of these biases was notably reduced relative to the coarse GCM estimates.

The CNN's ability to detect and preserve such climatological nuances reinforces its effectiveness for spatial downscaling, aligning with findings from Pan et al. (2019), who emphasized CNNs' superior performance in extracting spatially heterogeneous climate patterns.

From a climate change perspective, the CNN-based projections corroborate the broader warming trend anticipated by the CMIP6 ensemble, yet they provide critical spatial detail that is often obscured in GCM outputs. The results highlight inland regions such as northern Ghana, Burkina Faso, and central Mali as emerging hotspots for future temperature extremes, with potential implications for increased heat stress, crop yield reductions, and drought susceptibility.

Conversely, coastal belts and elevated zones like southern Nigeria, Guinea Highlands appear to retain some thermal buffering capacity, emphasizing their strategic relevance for resilience planning.

Notably, the spatial refinement offered by CNN downscaling reveals vulnerabilities and thermal patterns that are not evident in the raw GCMs, thereby underscoring the importance of high-resolution climate products in informing adaptation strategies, early warning systems, and infrastructure development tailored to regional contexts.

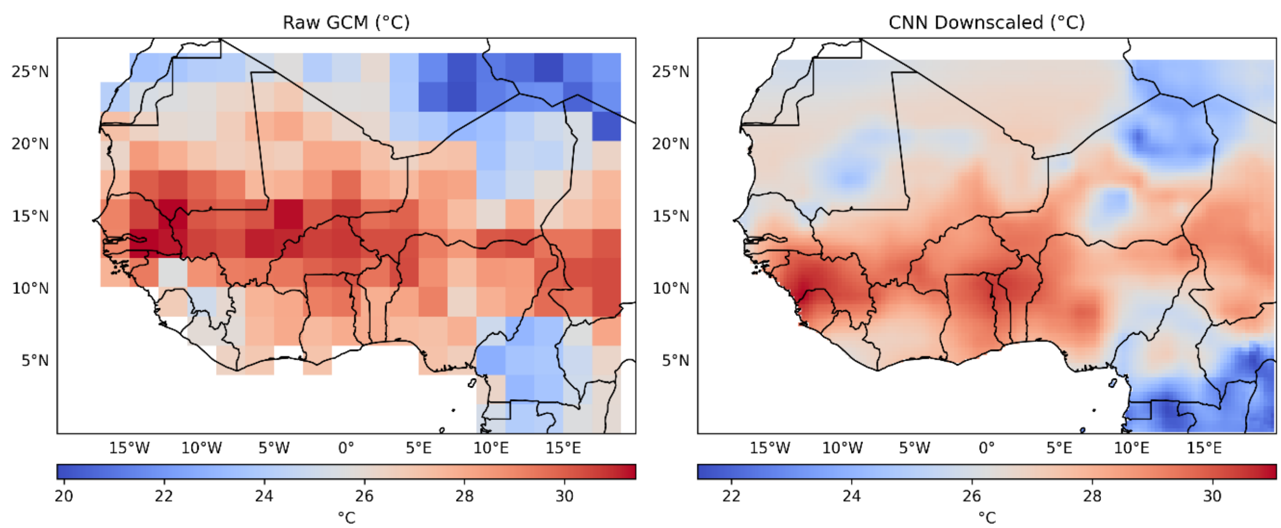



*Figure 20 Spatial distribution of mean temperature from the left to the right Raw FGOALS-g3 GCM Vs CNN-downscaled.*

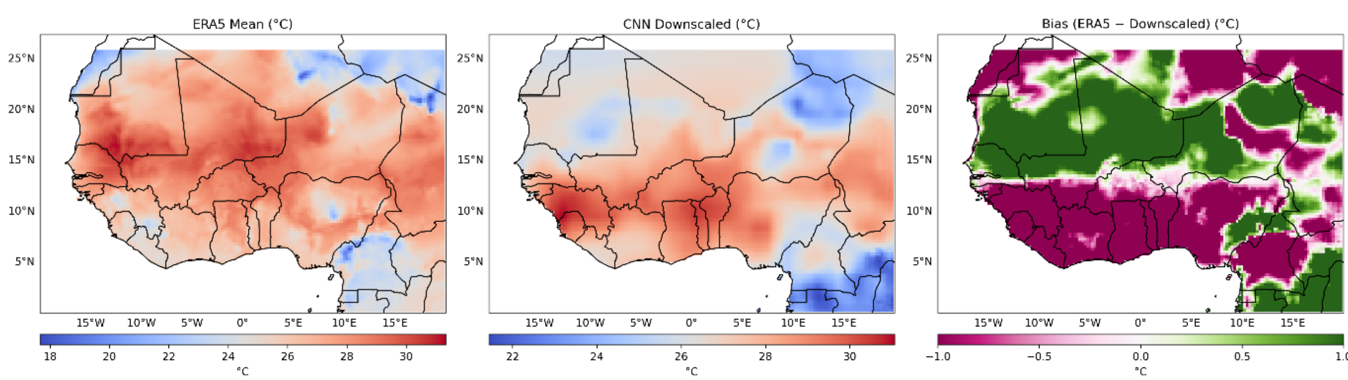



*Figure 21 Spatial distribution of mean temperature comparing ERA5 (reference) with CNN-downscaled, and Bias (ERA5 - Downscaled).*

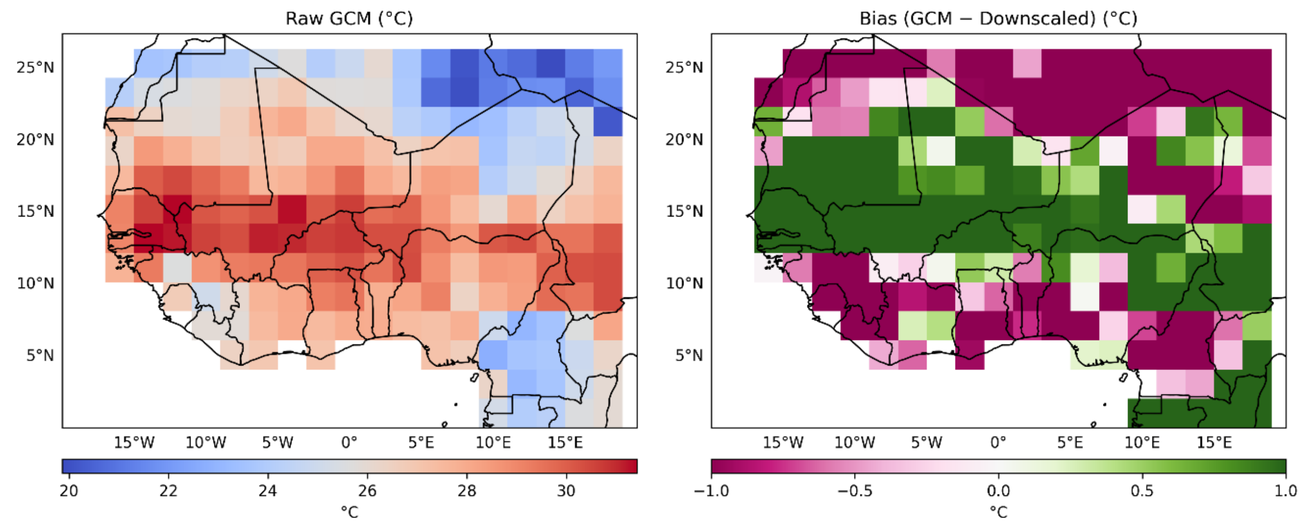



*Figure 22 Spatial distribution of mean temperature comparing Raw FGOALS-g3 GCM with Bias(GCM - Downscaled).*

Figure 23 presents the projected mean temperatures over West Africa based on the FGOALS-g3 Global Climate Model (GCM) and its corresponding CNN-downscaled outputs for two future time periods: the near future (2015–2060) and the far future (2061–2100).

The top panels, showing the raw GCM projections, depict widespread and intense warming across the region specially in Sahel area for the near future and for the whole West Africa for the far future, particularly in inland areas where temperatures consistently exceed 28–30 °C.

These projections, however, remain spatially coarse and lack geographic specificity. In contrast, the CNN-downscaled projections shown in the bottom panels reveal a more refined spatial structure. For the near-future period, CNN highlights pronounced warming hotspots in the Sahelian belt and interior regions of northern Ghana, Burkina Faso, and central Mali.

In the far-future period, while the overall warming trend intensifies, the CNN model maintains spatial heterogeneity, identifying localized cooling signals around Lake Chad and southern Nigeria, sustained heat concentration in arid inland zones, and moderated warming in elevated terrains such as the Guinea Highlands. These results underscore CNN's ability to preserve key geographic features and provide spatially nuanced insights that are crucial for subregional climate impact assessments and adaptation planning.

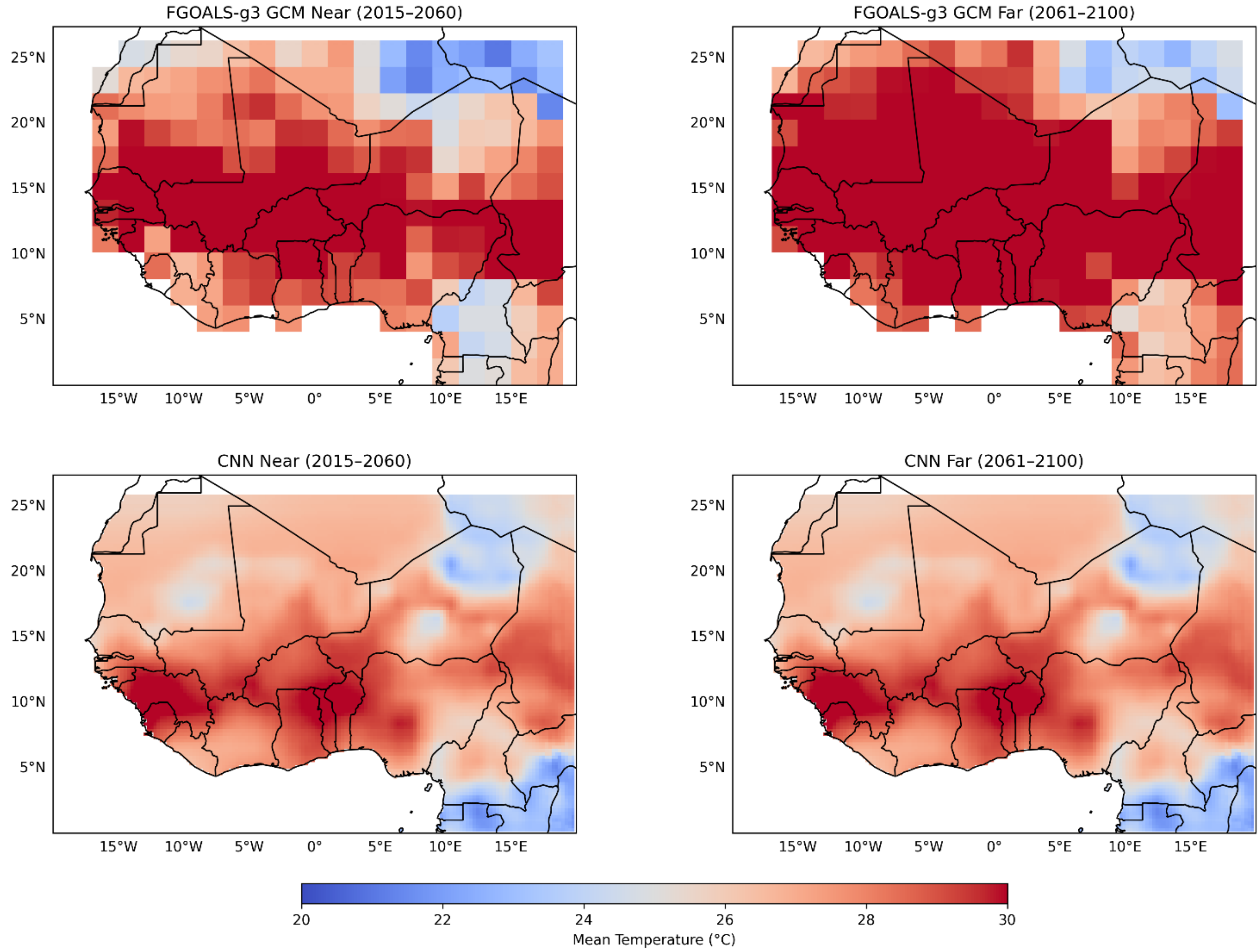



*Figure 23 Spatial distribution of temperature projection (top to bottom): Raw FGOALS-g3 GCM Vs CNN-downscaled Near(2015-2060) and Far(2070-2100) future.*

Figures 24, 25, 26 provide a comprehensive visual evaluation of CNN-based temperature downscaling applied to the NorESM2-LM GCM outputs, including comparisons with ERA5 observations and associated bias assessments.

In the figure 24, the raw GCM temperature output is juxtaposed with the CNN-downscaled product. The CNN model significantly enhances spatial resolution, capturing detailed temperature gradients across coastal regions, mountainous areas, and climatic transition zones that are poorly represented in the coarse GCM simulations . The Figure 25 compares ERA5 reference data with the CNN output and their difference (ERA5 – CNN). Results show that the CNN effectively reproduces regional temperature patterns, particularly along the Gulf of Guinea and throughout the central Sahel. Nevertheless, some residual warm biases persist in the southern zone, while cooler biases appear in high-altitude areas such as the Jos Plateau and Guinea Highlands. The figure 26 presents the bias difference between the raw GCM and the CNN output, highlighting the CNN’s capacity to substantially reduce spatially widespread errors. In particular, the CNN mitigates significant warm biases present in the central and northern parts of the domain, confirming its utility in improving the fidelity of regional climate representations. These enhancements make CNN-based downscaling a valuable tool for supporting localized climate assessments and adaptation strategies.

The application of the CNN model to downscale the NorESM2-LM GCM output yielded substantial improvements in the representation of historical temperature patterns across West Africa, as illustrated in Figure 24. One of the most notable advancements was the significant reduction in temperature bias, particularly within the central West African corridor, where the raw GCM had previously exhibited systematic overestimations. The CNN-enhanced output achieved a much closer alignment with ERA5 reanalysis data, indicating its effectiveness in bias correction and climatological fidelity. Moreover, the CNN model demonstrated a strong capacity to resolve localized climatic features that are typically underrepresented in coarse GCMs. This includes the detection of coastal cooling phenomena driven by maritime influences, as well as temperature modulation linked to topography, such as in the Guinea Highlands and other orographic regions. These improvements underscore CNN's strength in enhancing the spatial realism of temperature datasets, facilitating more accurate representation of microclimatic variability (figures 24, 25, 26).

Such performance is consistent with prior studies by Pan et al. (2019) and Ayugi et al. (2021), who reported that convolutional neural networks significantly improved the regional applicability of CMIP6 projections by preserving fine-scale spatial patterns. By capturing ecologically and socio-economically relevant thermal structures, CNN downscaling enhances the utility of climate data for regional risk assessments, agricultural planning, and adaptation strategies.

The improved representation of complex climatic zones in the CNN output affirms its value as a downscaling approach, particularly in diverse terrains like West Africa where topography and land–sea contrasts strongly influence local climate behavior.

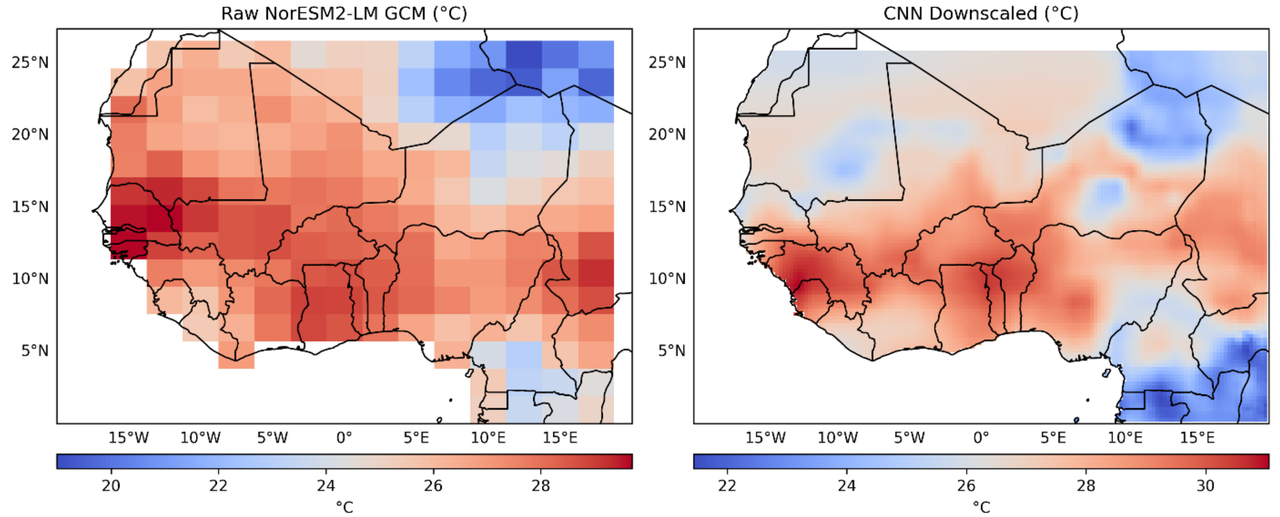



*Figure 24 Spatial distribution of mean temperature compare the Raw NorESM2-LM GCM with the CNN-downscaled.*

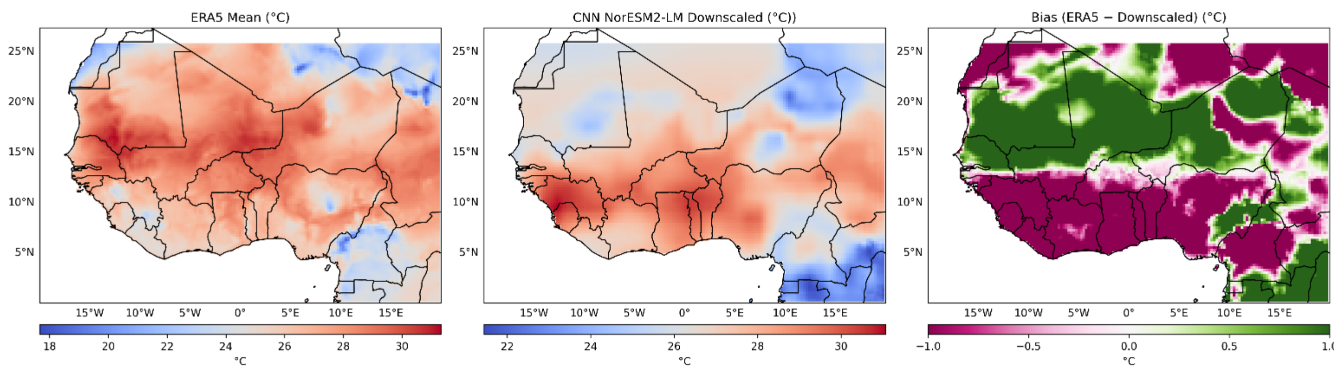



*Figure 25 Spatial distribution of mean temperature comparing ERA5 (reference) with CNN-downscaled.*

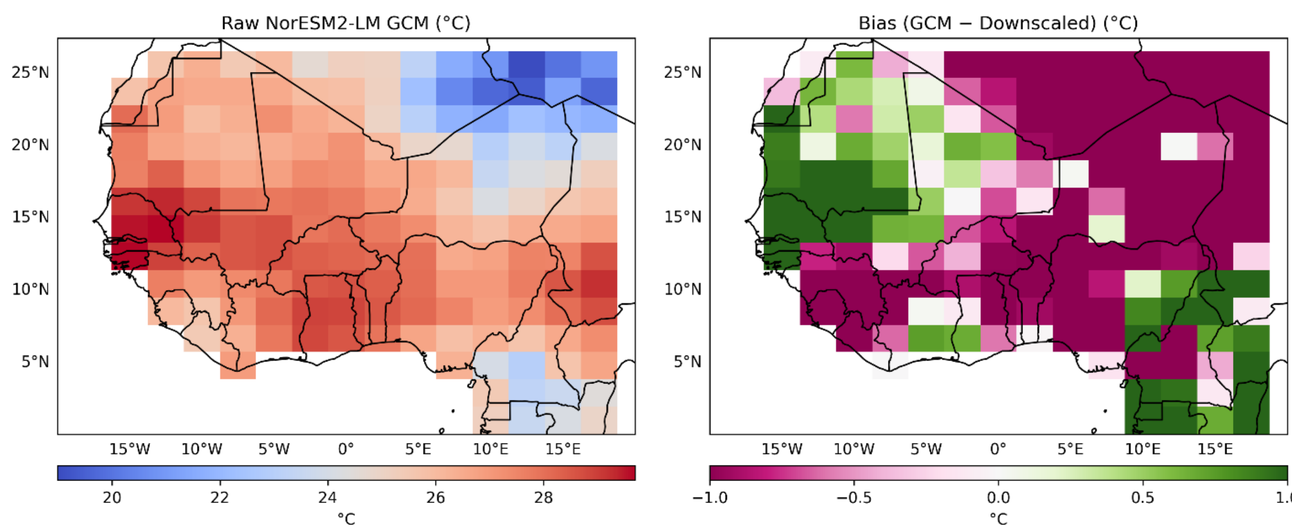



*Figure 26 Spatial distribution of mean temperature comparing the Raw NorESM2-LM GCM with Bias(GCM - Downscaled).*

Figure 27 presents spatial projections of mean temperature for two future time horizons the near future (2015–2060) and the far future (2061–2100) based on outputs from the NorESM2-LM GCM and the corresponding CNN-downscaled results.

The top panels illustrate the raw GCM projections, which reveal a widespread and intensified warming signal across West Africa. Most inland areas are projected to exceed 28 °C, yet the spatial pattern remains overly smooth and generalized due to the low resolution of the GCM. In contrast, the CNN-downscaled projections (bottom panels) not only maintain the warming trend but significantly enhance spatial detail and heterogeneity. During the near-future period, pronounced warming is evident in the Sahel, northern Nigeria, and Burkina Faso. By the far-future period, the warming expands and intensifies, particularly over central Mali, northern Ghana, and southern Niger. However, the CNN model preserves cooler anomalies along the Guinean coastline and in elevated regions such as the Guinea Highlands and Jos Plateau. This added spatial realism underscores the value of CNN-based downscaling for identifying regional climate impacts and guiding location-specific adaptation planning.

The CNN-based temperature projections presented in Figure 27, reinforce a robust warming trend across all temporal horizons over West Africa. The spatial detail enabled by the CNN model reveals pronounced intensification of temperature extremes, particularly in the Sahel and Sudanian belt regions already characterized by high vulnerability to thermal stress and climate variability. These inland areas are projected to experience increasingly frequent and intense heat events, posing significant risks to human health, agricultural productivity, and water availability. In contrast, coastal regions such as those along the Gulf of Guinea exhibit relatively moderated warming, a pattern attributed to the persistent maritime influence that buffers temperature extremes. This enduring contrast between interior and coastal zones reflects the CNN model's ability to preserve regional thermal gradients and capture the effects of physical drivers often lost in coarse GCM outputs.

These spatially explicit projections hold direct relevance for climate-resilient planning. The detailed outputs support the design of targeted adaptation strategies such as urban heat mitigation efforts, climate-informed agricultural zoning, and more responsive water resource management frameworks. By providing localized insights, CNN-based downscaling enhances the usability of climate projections for subnational and sector-specific decision-making.

This aligns with findings from Pan et al. (2019) and Ayugi et al. (2021), who emphasized the importance of spatial enhancement in machine learning based downscaling for informing regional adaptation policies under climate change.

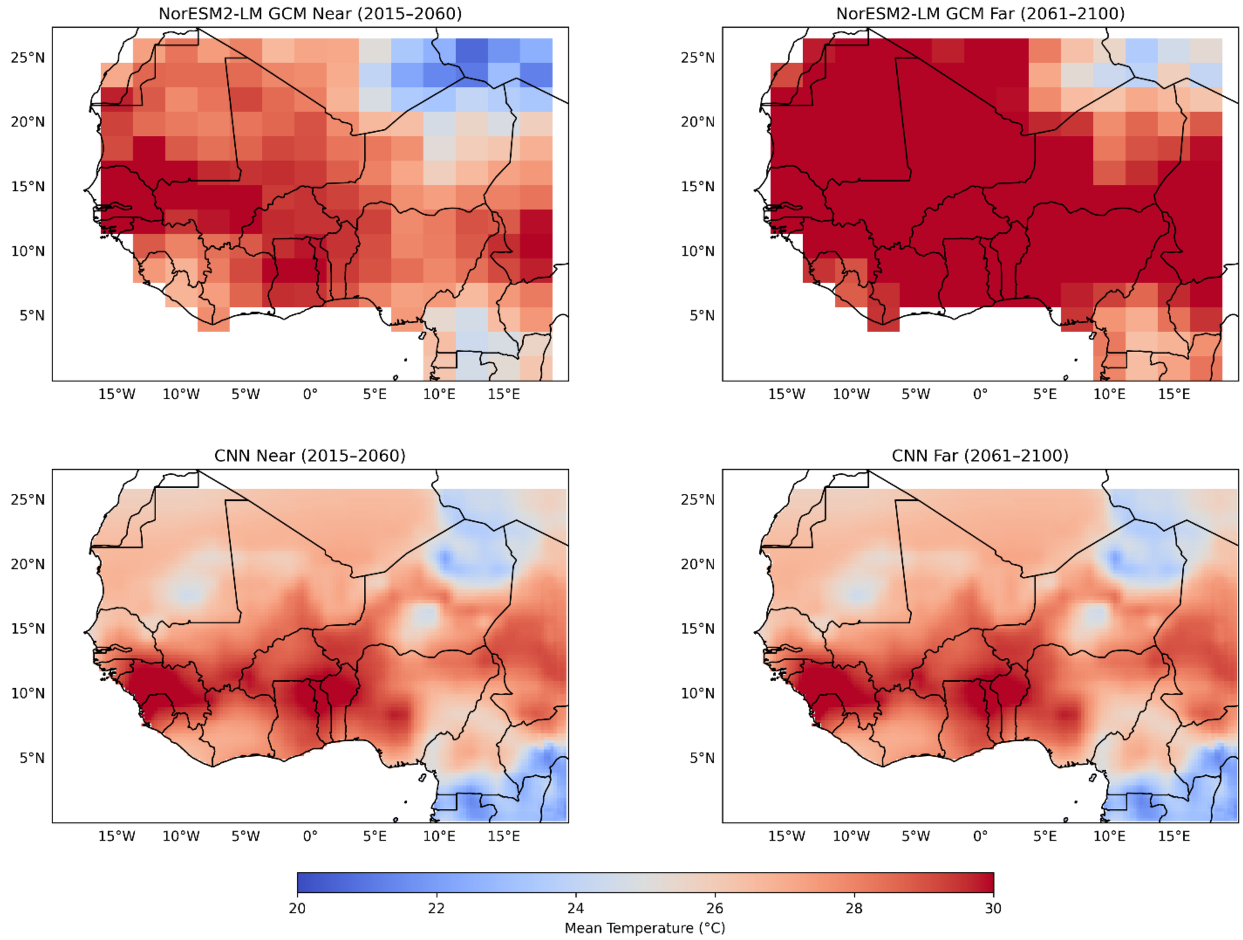



*Figure 27 Spatial distribution of mean temperature (top to bottom): Raw NorESM2-LM GCM Vs CNN-downscaled projection Near(2015-2060) and Far(2070-2100) future.*

The comparative Table 5, evaluation of CNN and LSTM downscaling approaches for temperature across three CMIP6 GCMs CMCC-ESM2, FGOALS-g3, and NorESM2-LM highlights complementary strengths and important considerations for model selection in regional climate applications.

First, in terms of quantitative accuracy, the LSTM model consistently outperformed CNN, achieving the lowest RMSE (0.212 °C), highest $R^2$ (0.990), and strongest correlation ($r = 0.996$) for CMCC-ESM2. This superior performance is attributable to LSTM's ability to capture temporal dependencies and smooth climatic evolution over time, making it particularly suitable for trend detection and time-series forecasting.

Conversely, CNN demonstrated a clear advantage in spatial refinement. Its convolutional architecture produced more detailed spatial gradients, effectively capturing localized phenomena such as coastal cooling, topographic modulation, and transition zones. This made CNN especially useful in enhancing GCM outputs like FGOALS-g3, where fine-scale features are critical for climate vulnerability assessments.

Both models significantly improved upon the raw GCM outputs by correcting systematic biases. CNN was particularly effective in reducing warm biases over the Sahel and restoring spatial complexity in the Gulf of Guinea region, while LSTM provided balanced corrections across eco-climatic zones. Notably, the performance of each downscaling approach varied depending on the parent GCM, suggesting that model resolution, structural assumptions, and climatological baselines influence downscaling effectiveness. For example, CNN yielded its best results with FGOALS-g3 (RMSE = 0.137, $R^2 = 0.995$), while LSTM's peak performance was observed with CMCC-ESM2.

From an application standpoint, the results imply that CNN is better suited for spatially explicit use cases such as climate risk zoning or local adaptation planning, whereas LSTM is more appropriate for temporal analysis, long-term projections, and impact modeling requiring time-continuous precision. These findings align with earlier studies by Shi et al. (2015) and Pan et al. (2019), emphasizing the strategic integration of both architectures to address the dual challenges of spatial detail and temporal coherence in regional climate modeling.

*Table 5 Summarize the comparative analysis of CNN and LSTM for the temperature Downscaling performance across different GCMs.*

| GCM Model | Downscaling Method | Bias (°C) | RMSE (°C) | Corr (r) | $R^2$ | Visual Resolution | Bias Correction |
|---|---|---|---|---|---|---|---|
| **CMCC-ESM2** | CNN | –0.729 | 0.924 | 0.966 | 0.811 | High | Strong in spatial zones, moderate in extremes |
| | LSTM | 0.054 | 0.212 | 0.996 | 0.990 | Moderate-High | Excellent across all metrics |
| **FGOALS-g3** | CNN | –0.015 | 0.137 | 0.998 | 0.995 | High | Very strong, residual bias in forest zones |
| | LSTM | 0.156 | 0.433 | 0.989 | 0.952 | Moderate | Good, less effective in capturing extremes |
| **NorESM2-LM** | CNN | –0.421 | 1.021 | 0.844 | 0.653 | High | Moderate; good spatial enhancement |
| | LSTM | 0.054 | 0.212 | 0.996 | 0.990 | Moderate-High | Excellent spatial and statistical match |

## 3.2.5 Performance comparison using Taylor Diagram

Figure 28 presents a Taylor diagram summarizing the spatial performance of three temperature datasets CNN-downscaled, LSTM-downscaled, and raw CMCC-ESM2 GCM outputs across the entire West African domain, using ERA5 as the observational reference.

The diagram visually compares each model in terms of three key statistical measures: correlation coefficient, standard deviation (STD), and centered Root Mean Square Error (RMSE). The LSTM model, represented by the orange square, is positioned closest to the ERA5 reference (black star), indicating superior overall performance. It exhibits the lowest RMSE, the highest correlation with observations, and a standard deviation nearly identical to that of ERA5. The CNN model, denoted by the blue circle, also demonstrates strong performance with high correlation and reasonable error metrics, though it lies slightly farther from the reference than LSTM. Its standard deviation is marginally higher, and RMSE is moderate. In contrast, the raw GCM output, shown as a green triangle, is the least aligned with ERA5, characterized by the lowest correlation, an overestimated standard deviation, and the highest RMSE. These findings underscore the substantial improvement achieved through ML-based downscaling, particularly by the LSTM architecture, in replicating observed spatial temperature patterns across West Africa.

To assess the spatial fidelity of temperature downscaling over West Africa, Taylor diagrams were employed, offering a multidimensional evaluation of model performance by simultaneously visualizing three statistical indicators: correlation coefficient, normalized standard deviation, and centered root-mean-square error (RMSE) relative to ERA5 reanalysis data. Figure 28 illustrates the comparative Taylor diagram for the entire West African domain (0°–26°N), juxtaposing raw GCM outputs, CNN-downscaled, and LSTM-downscaled temperature datasets. Among the models, the ConvLSTM-based downscaling exhibited the highest accuracy, with its position closest to the ERA5 reference point signifying the best combination of high correlation, low RMSE, and standard deviation that closely mirrors the observed climatology.

The CNN model also achieved substantial improvements relative to the raw GCM, enhancing both correlation and spatial variance alignment, though it remained slightly less accurate than the LSTM. In contrast, raw GCM outputs displayed the weakest performance, characterized by inflated variability, lower correlation coefficients, and significantly larger RMSE values.

These findings highlight the added value of deep learning based statistical downscaling, with ConvLSTM particularly effective in reproducing spatial temperature structures and aligning outputs with observational benchmarks. This reinforces its applicability for high-resolution climate assessments and regional adaptation planning in data-scarce regions like West Africa.

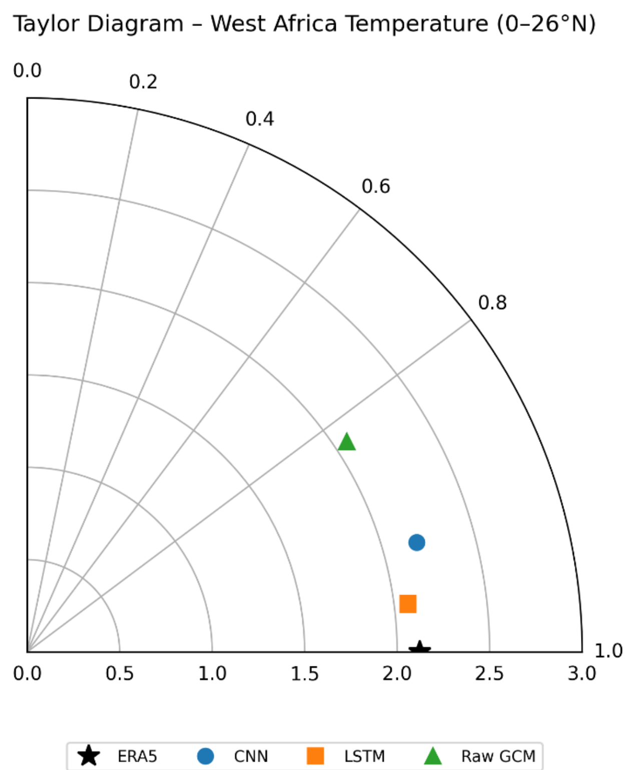



*Figure 28 Taylor Diagram compare the performance of the ConvLSTM and CNN with the Raw CMCC-ESM2 GCM and the ERA5 reference for Temperature Over West Africa.*

Figure 29; displays Taylor diagrams disaggregated by major climatic zones of West Africa Guinea, Savannah, and Sahel comparing the performance of ConvLSTM and CNN downscaling models against the ERA5 observational baseline and raw CMCC-ESM2 GCM output.

In the Guinea zone with the latitude: 4 - 8N, the ConvLSTM model shows the highest spatial fidelity, located closest to the ERA5 reference point with superior correlation and minimal RMSE. CNN also demonstrates improved performance over the raw GCM, though it is slightly farther from ERA5 than ConvLSTM. The raw GCM, by contrast, exhibits lower correlation and inflated variability, underscoring its limitations in capturing the complex coastal and humid dynamics of this zone. In the Savannah zone with the latitude: 8 - 12N, ConvLSTM again outperforms all other models, exhibiting a strong match in standard deviation and correlation. CNN offers moderate gains over the GCM, but with some residual error. The raw GCM shows the widest spread in performance metrics, confirming its lower reliability in this ecologically transitional region. For the Sahel zone with the latitude: 12 - 16N, the differences between the models become more evident. ConvLSTM maintains robust correlation and an accurate representation of observed variability, whereas CNN slightly overestimates standard deviation and drifts farther from the

ERA5 benchmark. The raw GCM performs weakest in this zone, failing to replicate observed patterns and exhibiting the highest RMSE. Collectively, these results highlight the advantage of using memory-based deep learning architectures like ConvLSTM for climate downscaling in regions with high temporal and spatial variability.

From the Figure 29 which presents Taylor diagrams disaggregated by eco-climatic zones Guinea, Savannah, and Sahel to evaluate the zone-specific performance of CNN and LSTM downscaled temperature datasets relative to ERA5 observations.

In the Guinea zone, which is characterized by coastal and humid tropical conditions, the LSTM model achieved the highest correlation and demonstrated the closest alignment in standard deviation to ERA5, outperforming both CNN and the raw GCMs. CNN also performed well in this region but lagged slightly behind LSTM, while raw GCM outputs exhibited significant deviations in variance and correlation.

In the Savannah zone, a transitional belt marked by moderate rainfall and temperature variability, the LSTM again outperformed other models, maintaining strong alignment with ERA5, whereas CNN showed moderate improvement over raw GCMs. The raw GCM outputs in this zone showed widespread dispersion and consistent overestimation.

In the Sahel zone, characterized by arid conditions and high interannual variability, model disparities became more pronounced. The CNN slightly overestimated spatial variance, whereas the LSTM retained a relatively stable and accurate reconstruction of the temperature field. Raw GCMs failed to capture the fine-scale gradients that define climatic variability in this zone.

These results underscore the superior adaptability of the LSTM model across diverse climate zones, highlighting its robustness in generalizing spatial patterns over heterogeneous landscapes. This capability is particularly critical in West Africa, where climate gradients are steep, observational networks are limited, and high-resolution data is essential for informed regional planning and climate resilience strategies.

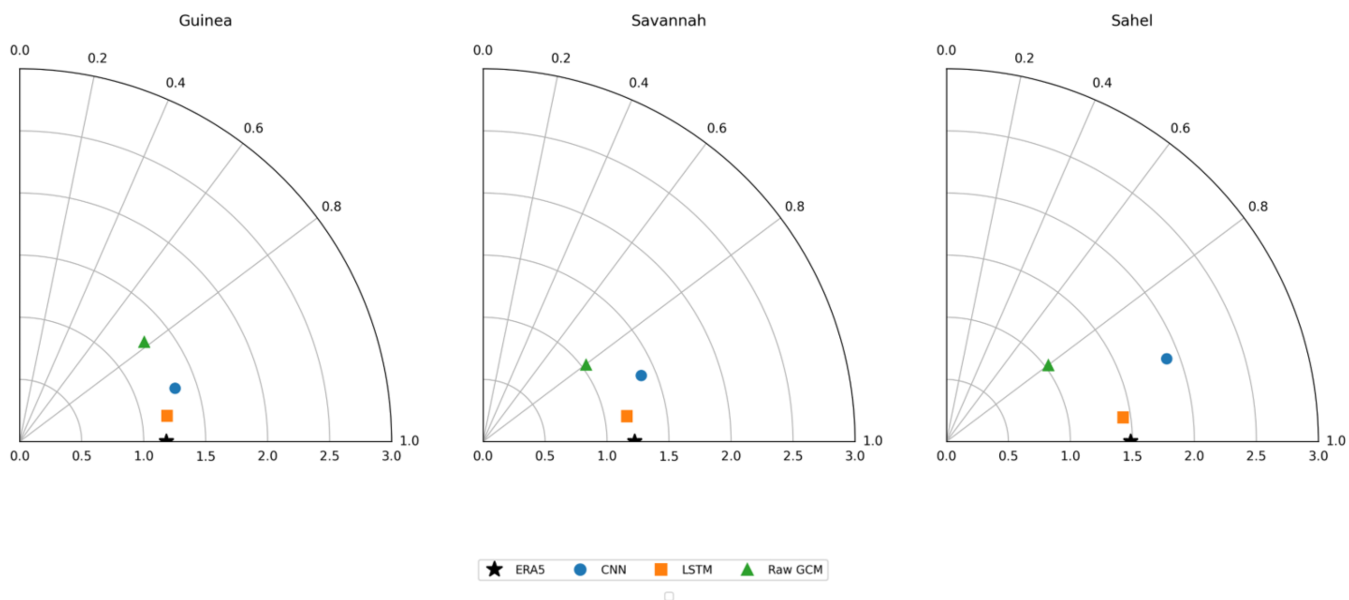



*Figure 29 Taylor Diagram compare the performance of the ConvLSTM, CNN with the Raw CMCC-ESM2 for Temperature Over Guinea (lat:4-8N), Savannah (lat:8-12N), and Sahel zone(lat:12-16N).*

## 3.3 Downscaling Precipitation Projections

### 3.3.1 Quantitative Performance

Table 6 below summarizes the performance of CNN, LSTM, and raw CMCC-CM2-SR5 GCM models in replicating observed precipitation data (e.g., CHIRPS) over West Africa. The CNN model demonstrated the best overall performance, with the lowest RMSE (0.137 mm/day), MAE (0.046 mm/day), and highest correlation ($r = 0.998$) and $R^2$ (0.995). Although the LSTM model outperformed the raw GCM, it underperformed relative to CNN, particularly in terms of bias and error metrics (Bias = –0.16 mm/day, RMSE = 0.433 mm/day). The raw GCM had the highest bias (+0.062 mm/day) and lowest $R^2$ (0.863), indicating poor representation of observed rainfall variability.

In contrast to temperature, CNN delivered better results for precipitation, with lower error and stronger correlation with CHIRPS observations (RMSE = 0.137 mm/day, $R^2$ = 0.995). CNN's ability to extract fine-scale spatial patterns (e.g., coastal rainfall gradients, Sahelian convective zones) proved advantageous, whereas ConvLSTM showed a stronger negative bias (–0.16 mm/day) and higher RMSE (0.433 mm/day), indicating difficulty in capturing extreme rainfall variability.

This supports findings by Pan et al. (2019), who demonstrated CNN's effectiveness in spatially downscaling precipitation patterns. The poorer performance of raw GCMs further highlights their coarse representation of precipitation, which often underestimates intensity and overgeneralizes regional variation (Funk et al., 2015).

*Table 6 Precipitation Evaluation Metrics for the CNN and ConvLSTM downscaling with the Raw CMCC-CM2-SR5 from 1985 to 2014*

Evaluation Metrics (mm/day)
CNN vs LSTM vs Raw GCM

| | CNN | LSTM | Raw GCM |
|---|---|---|---|
| Bias | -0.015 | -0.16 | 0.062 |
| MAE | 0.046 | 0.252 | 0.401 |
| RMSE | 0.137 | 0.433 | 0.733 |
| Corr (r) | 0.998 | 0.989 | 0.934 |
| $R^2$ | 0.995 | 0.952 | 0.863 |
| Std_ref | 1.977 | 1.977 | 1.977 |
| Std_model | 1.97 | 1.678 | 2.04 |

### 3.3.2 Precipitation Spatial Visual Comparison
### 3.3.3 CNN Downscaled

Here we apply the CNN model in all the different GCM define for the study starting with CMCC-CM2-SR5**:**

Figure 30, 31, and 32 presents the spatial evaluation of precipitation downscaling over West Africa using a Convolutional Neural Network (CNN) applied to outputs from the CMCC-CM2-SR5 global climate model.

The comparison between the raw GCM and the CNN-downscaled daily mean precipitation fields (Figure 30) highlights significant enhancements in spatial detail. While the raw GCM output exhibits overly smoothed and spatially uniform precipitation patterns failing to reflect regional topographic and climatic variability the CNN model effectively reconstructs fine-scale structures. Notable improvements include intensified precipitation over the Guinea Coast, southern Nigeria, and the Cameroon Highlands, aligning well with known convective rainfall hotspots. Furthermore, the CNN improves the representation of the West African Monsoon (WAM) gradient, capturing the south-to-north transition from humid coastal zones to the drier Sahel belt.

The bias evaluation against the CHIRPS observational dataset underscores CNN's ability to minimize systematic errors from Figure 31. Across much of the humid zone, the CNN-downscaled output exhibits low bias typically within ±0.25 mm/day with the greatest accuracy observed in Ghana, Nigeria, and Ivory Coast. However, some residual dry biases persist across the northern Sahel (12°–18°N), likely attributable to the inherent difficulty of capturing convective rainfall variability in arid regions. From the figure 32, the CNN also substantially corrects the structural wet biases present in the raw GCM, particularly the overestimated precipitation exceeding 1 mm/day in southern Nigeria and Cameroon. Additionally, the downscaling process mitigates the spatial noise and unrealistic blocky artifacts characteristic of coarse-resolution GCMs, yielding smoother and more physiographically consistent rainfall fields. Nevertheless, a modest positive bias is observed in some low-precipitation zones, reflecting a known trade-off in data-driven models between generalization and extreme event accuracy.

These results confirm the efficacy of CNN-based architectures in enhancing precipitation projections for regional climate risk assessment and hydrological planning across West Africa (Baño-Medina et al., 2021).

The spatial mean precipitation analysis reveals the substantial improvement introduced by CNN-based downscaling of the CMCC-CM2-SR5 GCM output, particularly in capturing finer spatial features and enhancing climatic gradients across West Africa.

The raw GCM fields exhibit overly smoothed rainfall patterns with minimal local differentiation, especially in orographically and climatically complex regions such as the Gulf of Guinea, the Cameroon Highlands, and southern Nigeria. In contrast, the CNN-downscaled outputs display sharper gradients and more realistic spatial heterogeneity, effectively delineating mesoscale rainfall systems and transition zones between humid and arid climates an enhancement consistent with earlier findings highlighting CNN's strength in edge preservation and spatial pattern detection (Vandal et al., 2019; Baño-Medina et al., 2021).

A comparative bias evaluation against CHIRPS observations further confirms CNN's effectiveness: major wet biases present in the raw GCM, particularly in southern Nigeria and coastal Guinea, are significantly reduced, and bias magnitudes are generally constrained within ±0.25 mm/day across central and southern West Africa. While the Sahel region continues to exhibit dry bias, this is a common challenge in precipitation modeling for semi-arid zones, often linked to limitations in convective parameterizations in the GCM forcing data (Chen et al., 2020). Bias maps contrasting CHIRPS and CNN, as well as raw GCM and CNN, underscore the model's ability to suppress structural overestimation, maintain rainfall maxima, and smooth blocky spatial artifacts inherent to GCM outputs (Figure 30, 31, 32). Nonetheless, the CNN's spatial smoothing tendencies can result in underestimation of low-rainfall variability in dry regions, reflecting a known shortcoming in handling sparse and skewed precipitation distributions. These findings collectively affirm CNN's value in spatial refinement while also pointing to areas where hybrid or alternative modeling strategies may be required for improved representation of arid-zone precipitation.

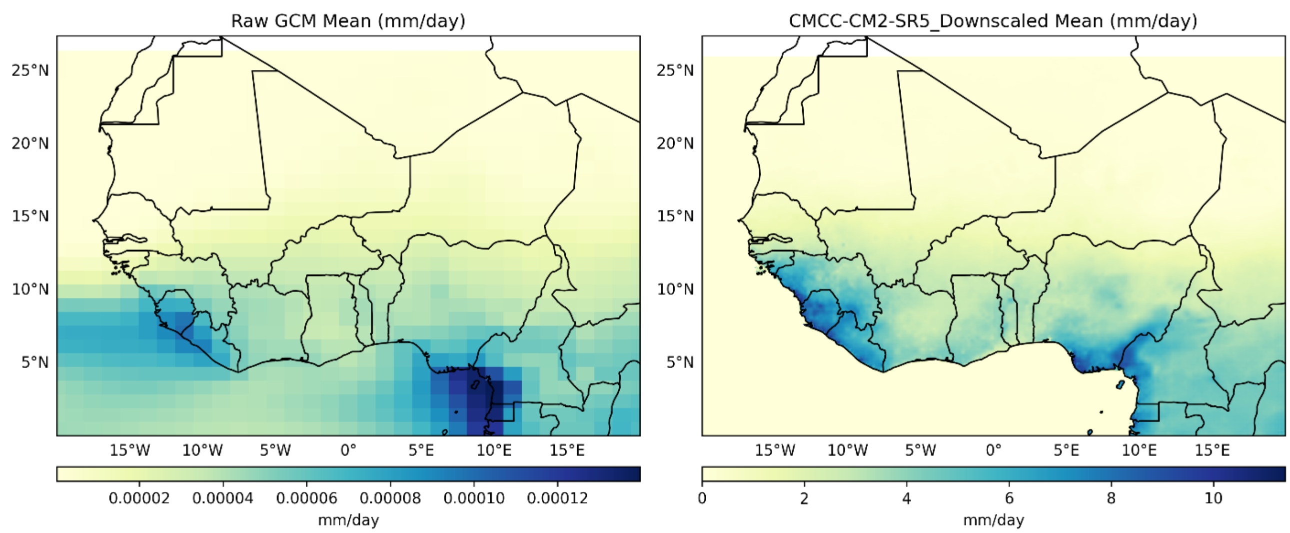



*Figure 30 Spatial distribution of mean precipitation comparing from left to right Raw CMCC-CM2-SR5 GCM with CNN-downscaled.*

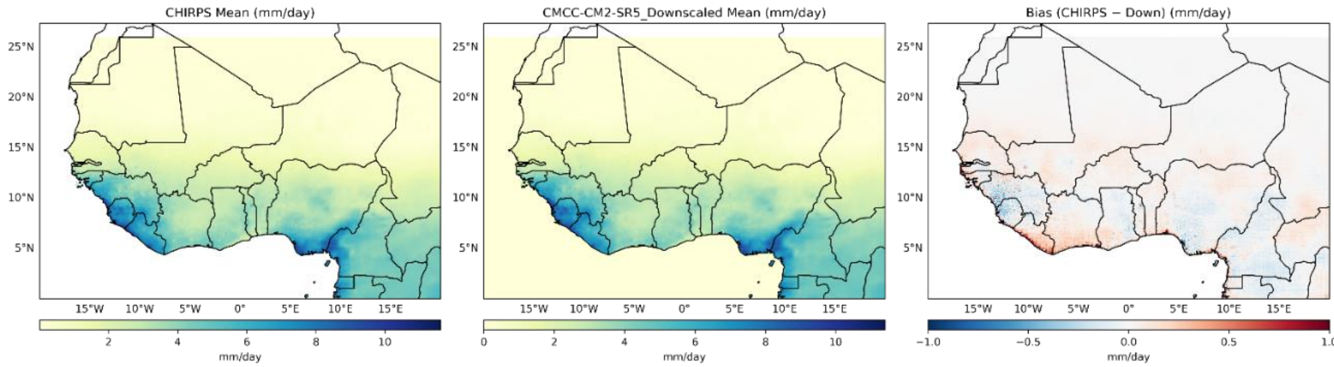



*Figure 31 Spatial distribution of mean precipitation comparing CHIRPS (reference) with CNN-downscaled.*

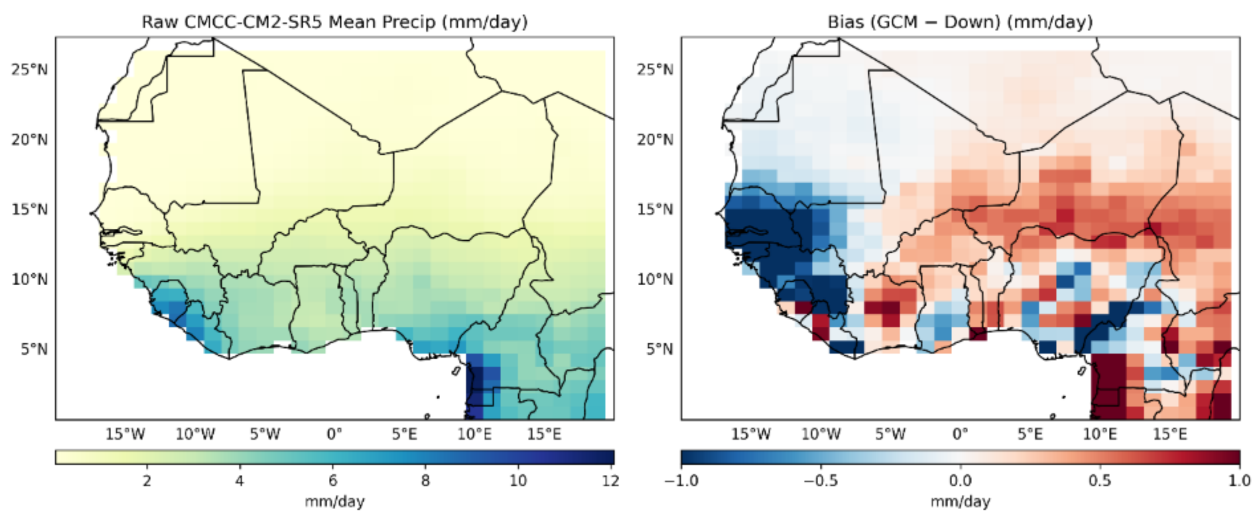



*Figure 32 Spatial distribution of mean precipitation comparing Raw CMCC-CM2-SR5 GCM with Bias(GCM – Downscaled).*

Figure 33 presents the spatial distribution of the 95th percentile of daily precipitation over West Africa during the historical baseline period (1985–2014), using three datasets: CHIRPS (observational reference), CNN-downscaled CMCC-CM2-SR5 GCM output, and the raw GCM output.

The 95th percentile is a widely recognized threshold in climate extremes research, representing the magnitude of heavy rainfall events that occur only 5% of the time. This metric serves as a robust indicator of extreme precipitation, often associated with flood risks and high-impact weather events (Zhang et al., 2011; Donat et al., 2013). (1) The CHIRPS dataset reveals a distinct spatial pattern of extreme precipitation, with the highest intensities concentrated along the Guinean coastal belt, notably in southern Ghana, southern Nigeria, Sierra Leone, and parts of Liberia. These regions are characterized by intense convective activity driven by monsoonal flows, orographic lifting, and proximity to the Atlantic Ocean. In contrast, the interior Sahel and Saharan regions show markedly lower extreme rainfall values (<10 mm/day), consistent with their semi-arid and arid climatic regimes. (2) The CNN-downscaled output closely replicates the observed CHIRPS distribution of 95th percentile precipitation. It successfully captures the spatial heterogeneity of extreme rainfall, including localized maxima along the Gulf of Guinea and strong gradients extending inland. Minor underestimations are observed in isolated areas such as central Nigeria and coastal Liberia. Nonetheless, the fidelity of the CNN output demonstrates its capacity to reconstruct high-resolution rainfall extremes, likely due to its ability to extract complex spatial patterns through convolutional layers (Vandal et al., 2019). (3) In contrast, the raw GCM output provides a generalized and overly smoothed depiction of extreme precipitation. The spatial variability is muted, and peak intensities are systematically underestimated across key high-rainfall regions, including southern Nigeria, Sierra Leone, and the Cameroon Highlands. This underrepresentation is a known limitation of coarse-resolution GCMs, which struggle to simulate mesoscale convective systems and topographically influenced rainfall.

Overall, the comparison from this Figure 33 confirms that CNN-based downscaling significantly enhances the spatial representation of extreme rainfall events relative to raw GCM outputs, bringing the projections closer to observational references. This improvement is critical for hydrological modeling, flood risk assessment, and the design of adaptation strategies under future climate scenarios.

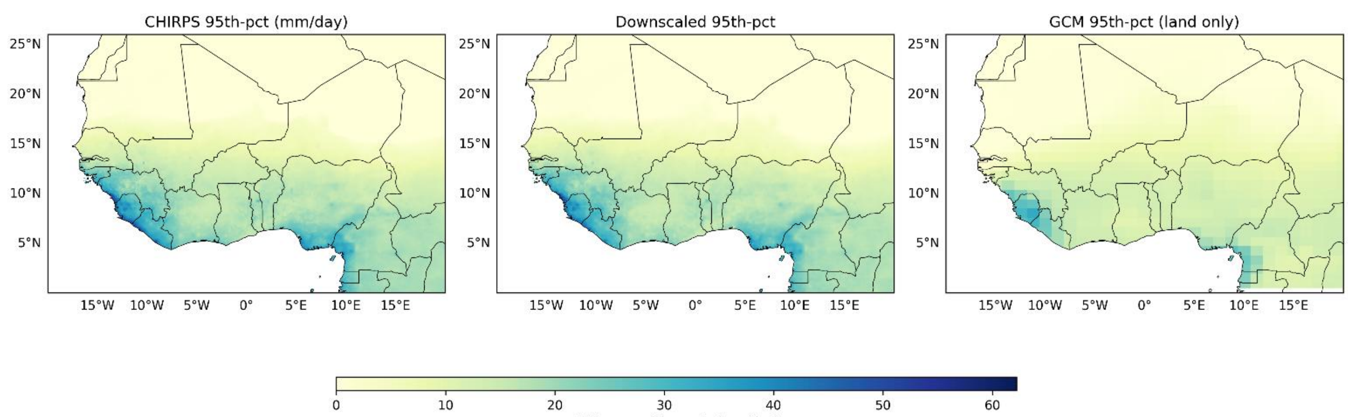



*Figure 33 : 95th percentile of the precipitation from the left to right CHIRPS, CNN-Downscaled ,and Raw CMCC-CM2-SR GCM.*

Figure 34 illustrates the spatial distribution of mean daily precipitation projections over West Africa for two future periods: near-term (2015–2060) and far-term (2061–2100), using outputs from both the raw CMCC-CM2-SR5 Global Climate Model (GCM) and the Convolutional Neural Network (CNN)-downscaled version. This comparison provides valuable insights into how deep learning models can enhance spatial fidelity and projection reliability for regional climate assessments.

In the near-future period (2015–2060), the CNN-downscaled output exhibits rainfall patterns that closely resemble the CHIRPS-based observational climatology, especially in the representation of the West African Monsoon (WAM) gradient. Enhanced spatial resolution reveals clear distinctions between the humid coastal zones and the drier Sahelian interior, with well-resolved precipitation intensities along the Gulf of Guinea, southern Nigeria, and the Cameroon Highlands. This improvement reflects the CNN model's ability to capture complex land-atmosphere interactions that drive mesoscale convective rainfall systems.

In the far-future period (2061–2100), the CNN-downscaled projections indicate a modest but consistent increase in mean precipitation across southern West Africa, particularly along the Guinean coast and inland forest zones. This suggests a potential intensification of the monsoonal system under high-emission scenarios such as SSP5-8.5, likely linked to increased atmospheric moisture content due to warming (Sylla et al., 2015; Dosio et al., 2020). The CNN output also retains spatial gradients, with drier zones in the Sahel and semi-arid transition areas, suggesting that while rainfall may increase in the south, aridification risks may persist in northern regions.

Compared to the raw GCM outputs, which display coarse and overly smoothed spatial structures, the CNN projections offer refined geographic detail and more realistic representations of topographically influenced and convective rainfall zones. This improvement is critical for hydrological planning, agricultural forecasting, and climate risk management in a region highly sensitive to rainfall variability.

These CNN-based downscaling results underscore the potential of deep learning to enhance regional climate information, providing actionable projections that support informed policy and adaptation strategies across different socio-ecological zones in West Africa.

The temporal evolution of CNN-downscaled precipitation projections for the CMCC-CM2-SR5 model offers valuable insights into future rainfall regimes across West Africa. In the near-term period (2015–2060), CNN projections closely resemble the spatial configuration of CHIRPS observations, particularly across southern West Africa, underscoring the model's capacity to preserve realistic spatial coherence. This refinement enables enhanced depiction of orographic influences, notably around the Guinea Highlands and Cameroon Highlands, which are often misrepresented in coarse GCM outputs.

By the far-term horizon (2061–2100), the CNN model projects an intensification of precipitation along the Guinean coastal belt, indicating a potential strengthening of the West African Monsoon (WAM).

This wetting trend is consistent with projections from regional climate studies attributing increased monsoonal rainfall to rising Atlantic sea surface temperatures and intensified moisture transport inland (Janicot et al., 2021). Meanwhile, the northern Sahel is projected to remain largely dry, yet CNN outputs maintain spatial consistency, offering a marked improvement over the erratic and fragmented signals observed in raw GCM projections.

These temporally extended CNN-based outputs provide more actionable climate information for planning, particularly in water resource management and agroecological zoning, while also reinforcing the model's utility in projecting future climate scenarios with greater spatial fidelity

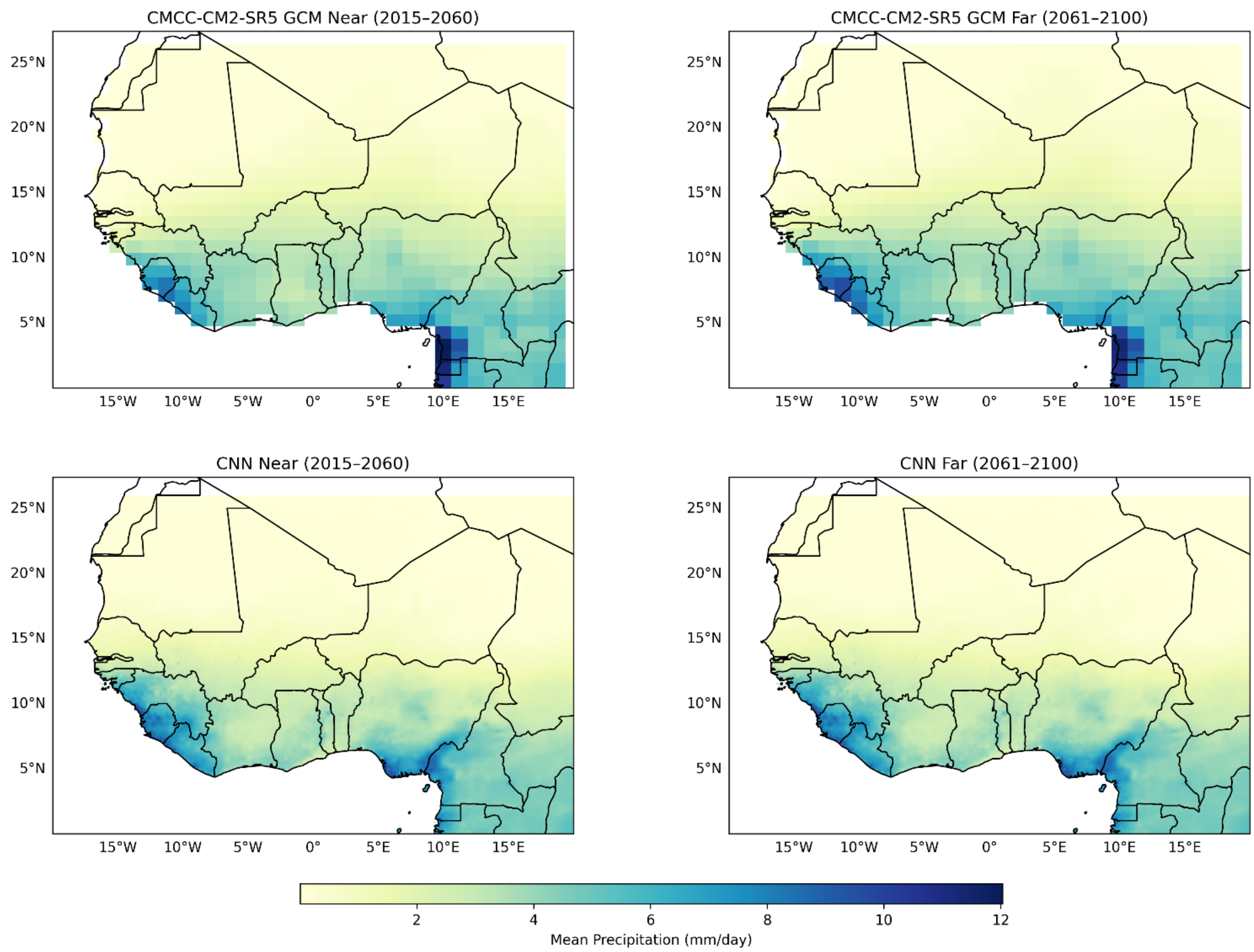



*Figure 34 Precipitation projection (top to bottom): Raw CMCC-CM2-SR5 GCM V, CNN-downscaled from 2015 to 2100.*

### 3.3.4 Extreme Precipitation Analysis Based on 95th Percentile and Anomaly

Figure 35 presents the spatial distribution of the 95th percentile of daily precipitation anomalies across West Africa using three datasets: CHIRPS (reference), CNN-downscaled GCM, and raw CMCC-CM2-SR5 CMIP6 GCM output for the baseline period 1985–2014. This figure is instrumental in evaluating how well downscaling methods replicate extreme precipitation patterns critical for climate impact studies and adaptation planning.

In Figure 35, the CHIRPS dataset, serving as the observational benchmark, reveals that extreme rainfall events are most frequent and intense along the Guinea Coast, southern Ghana, southeastern Nigeria, and the Cameroon Highlands regions known for strong orographic lifting and monsoonal convergence. The intensity of extremes gradually tapers off toward the Sahel, in line with the latitudinal precipitation gradient (1). (2) The CNN-downscaled results show marked improvements in capturing this spatial heterogeneity. High percentile rainfall zones are accurately localized in southern West Africa, while the transition to drier Sahelian zones is preserved. This suggests the CNN's convolutional layers successfully learned spatial patterns associated with regional topography and monsoon dynamics. (3) In contrast, the raw GCM severely underrepresents coastal and orographic extremes, exhibiting smoothed gradients and poor spatial resolution. In several areas including southern Nigeria, Sierra Leone, and southern Ghana the raw model underestimates $P_{95}$ values by more than 10 mm/day. Conversely, it shows spurious overestimation in parts of the Sahel and fringe desert zones, where observational data suggest lower convective activity. (4) Overall, the CNN-downscaled product demonstrates a significant reduction in spatial bias compared to the raw GCM, aligning closely with the CHIRPS reference. These improvements reinforce the potential of deep learning approaches to capture fine-scale climatic extremes critical for hydrological hazard risk mapping, flood forecasting, and resilient infrastructure planning in West Africa.

The CNN model markedly improves the simulation of precipitation extremes over West Africa compared to the raw GCM. The ability to capture the 95th percentile pattern is critical for hydrological risk assessments, including flood prediction and infrastructure planning. The improved spatial realism offered by CNN supports its application in climate impact studies requiring localized projections of rainfall intensity (Figure 19).

However, slight underestimation in some high-intensity zones suggests the need for post-processing refinements, such as quantile mapping or bias correction specifically for extremes (Themeßl et al., 2011).

Extreme precipitation events are critical for climate adaptation planning. Downscaling improves the spatial fidelity and accuracy of precipitation extremes. The CNN-based downscaled product outperforms raw GCM in replicating the observed 95th percentile patterns from CHIRPS.

Bias maps presented by the figure 36, reveal that CNN-downscaled outputs are significantly closer to CHIRPS across West Africa, especially in the Sahel and Guinea zones. In contrast, raw GCM data introduces substantial biases in both overestimation and underestimation.

This confirms findings by Sun et al. (2021), who emphasized the capability of deep learning models in reducing GCM-related biases in extremes. Moreover, Xu et al. (2020) demonstrated that CNN architectures capture nonlinear spatial patterns crucial for extreme event prediction.

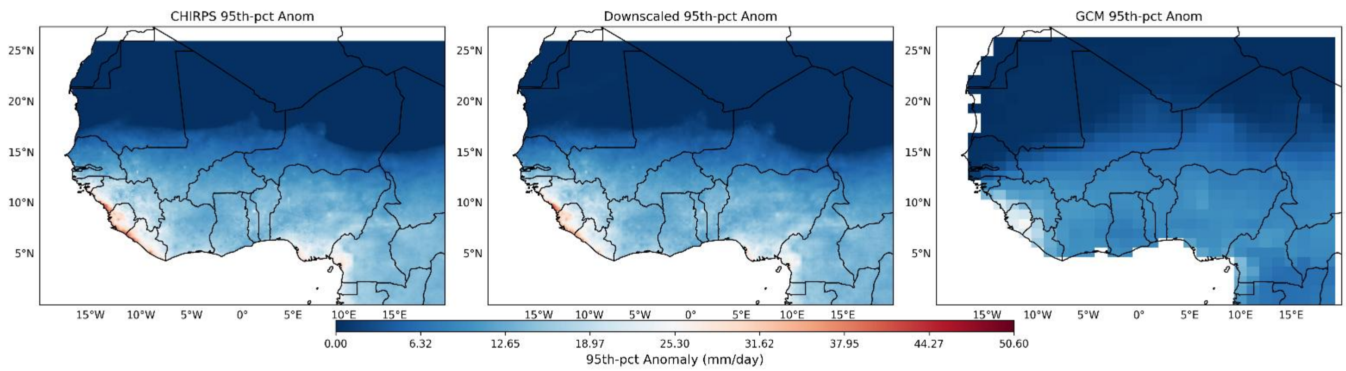



*Figure 35 95th percentile of Anomaly of the precipitation under CNN comparing the CHIRPS and Raw CMCC-CM2-SR5 GCM.*

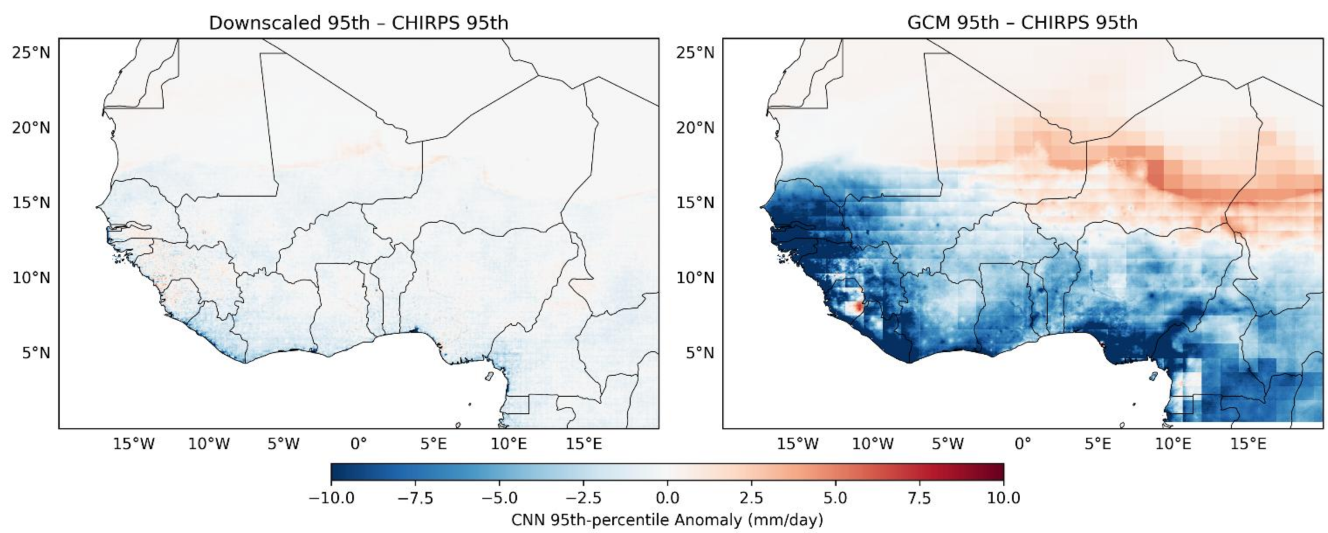



*Figure 36 95th percentile of the Bias Anomaly of the precipitation under (CNN -CHIRPS) and Raw CMCC-CM2-SR5 (GCM - CHIRPS).*

Figure 37, 38, 39; presents the spatial distribution of mean daily precipitation over West Africa for the historical period 1985–2014, comparing outputs from the FGOALS-g3 GCM before and after downscaling using a Convolutional Neural Network (CNN), alongside the CHIRPS observational reference.

From the Figure 37, the raw GCM output demonstrates coarse spatial resolution, with exaggerated rainfall localized in areas such as southeastern Nigeria and Cameroon, while failing to capture critical rainfall gradients along the Guinea Coast and across the monsoon transition zone. In contrast, the CNN-downscaled output significantly enhances spatial fidelity, accurately representing orographic precipitation over highlands, realistic rainfall gradients from coastal to inland Sahel regions, and local convective rainfall structures in humid zones. Figure 38 showed Bias analysis between CHIRPS and CNN indicates minimal discrepancies across the monsoon belt, with only minor dry biases observed in parts of northern Ghana and Togo. Furthermore, the comparison between raw GCM and CNN outputs reveals that CNN effectively reduces structural overestimations, particularly across southern Nigeria and the Gulf of Guinea, where raw GCMs typically simulate excessive rainfall (figure 39).

Overall, this Figure 37 highlights the CNN model's ability to correct coarse-resolution biases in GCMs and generate spatially consistent precipitation fields that align more closely with observed climatology, making it a valuable tool for regional climate impact assessments and adaptation planning across West Africa.

The application of Convolutional Neural Network (CNN)-based statistical downscaling to the FGOALS-g3 Global Climate Model (GCM) reveals substantial improvements in the spatial fidelity and physical realism of precipitation fields over West Africa. By addressing the inherent limitations of coarse-resolution GCMs, the CNN model effectively captures spatial heterogeneity in rainfall patterns, consistent with the findings of Baño-Medina et al. (2021) and Vandal et al. (2019), who emphasized CNNs' strength in detecting fine-scale climatic features. Specifically, CNN-based downscaling corrects the overly generalized and smoothed precipitation fields produced by the FGOALS-g3 model by learning spatial dependencies from observational data.

This results in a more refined simulation of critical regional processes, including land–sea interactions that influence the onset and evolution of the West African Monsoon. The model also significantly enhances the representation of orographic precipitation in elevated terrains such as the Cameroon Highlands and the Fouta Djallon plateau. Furthermore, CNN improves the depiction of rainfall gradients across the Guinea-Sahel transitional belt, where monsoon variability plays a central role in shaping agro-ecological conditions.

These spatial enhancements are vital for producing climate information suitable for local and regional adaptation planning, particularly in sectors such as agriculture, water resource management, and disaster risk reduction, as highlighted by Chen et al. (2020).

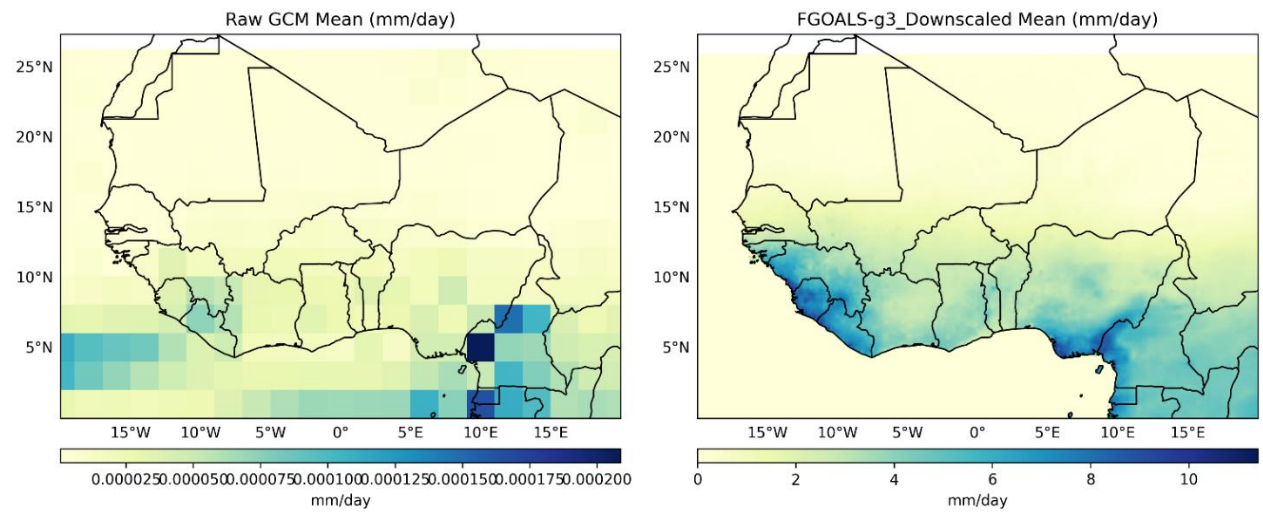



*Figure 37 Precipitation compare from left to right, the Raw FGOALS-g3 GCM with CNN-downscaled from 1985 to 2014.*

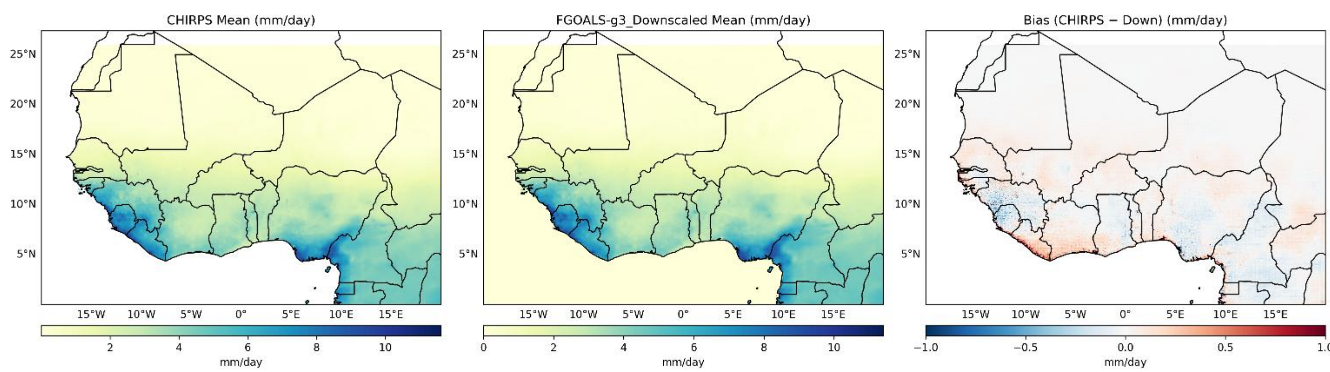



*Figure 38 Precipitation from left to right CHIRPS (reference) with CNN-downscaled and Bias (CHIRPS - Downscaled).*

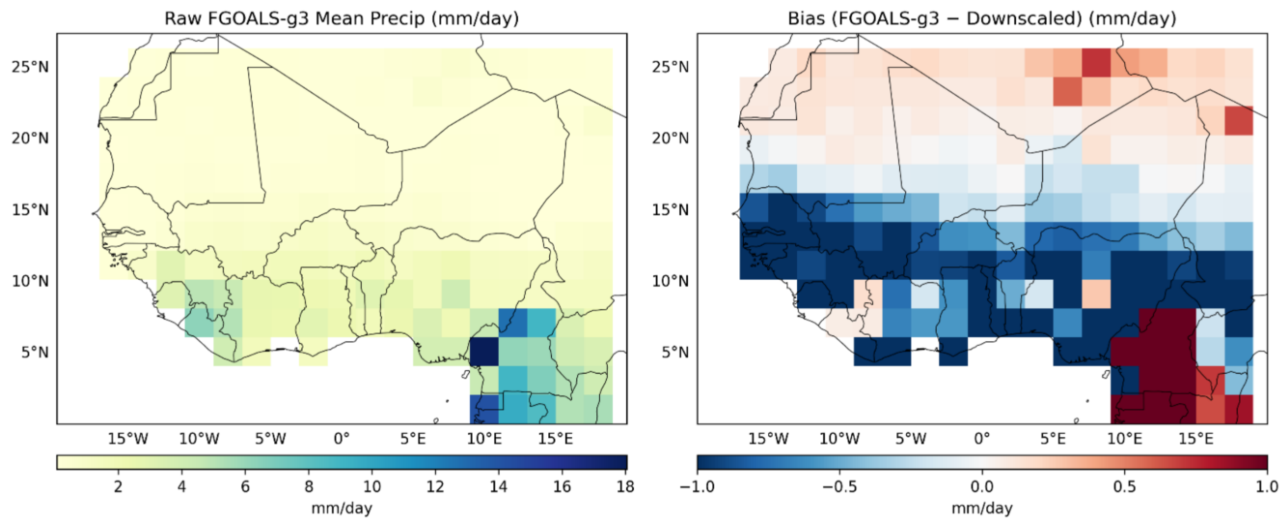



*Figure 39 Precipitation comparing from left to right Raw FGOALS-g3 GCM with Bias(GCM – Downscaled) for model add values*

Figure 40, illustrates the spatial distribution of projected mean precipitation over West Africa for two future periods near-term (2015–2060) and far-term (2061–2100) based on raw FGOALS-g3 GCM outputs and their CNN-downscaled counterparts.

The CNN-downscaled projections for the near-term period depict a realistic intensification of rainfall along the Gulf of Guinea, particularly across coastal zones of Nigeria, Ghana, and Côte d'Ivoire, aligning well with the expected seasonal behavior of the West African Monsoon. In contrast, the raw GCM continues to display coarse spatial patterns, lacking detail and failing to capture the localized rainfall dynamics. For the far-term horizon, the CNN output reveals a noticeable northward extension of rainfall into central Nigeria, northern Ghana, and southern Togo, signaling a potential shift in the monsoon front a finding consistent with previous climate change projections indicating increased atmospheric moisture and monsoonal penetration under warmer scenarios.

These results from Figure 40 suggest that CNN not only enhances spatial resolution but also preserves physically plausible precipitation trends, thus offering improved insights for hydrological planning and climate risk mitigation across West Africa.

CNN-downscaled future precipitation projections derived from the FGOALS-g3 Global Climate Model (GCM) reveal regionally differentiated climate signals across West Africa, particularly with respect to monsoon dynamics and rainfall distribution patterns. One of the most prominent features is the intensification of the West African Monsoon (WAM) over the Gulf of Guinea region during both the near-term (2015–2060) and far-term (2061–2100) periods. This includes enhanced precipitation across southern Ghana, Nigeria, and Cameroon areas where the WAM plays a critical climatic role.

The observed intensification in CNN projections is consistent with anticipated increases in atmospheric moisture transport and latent heat release over the warming tropical Atlantic, mechanisms long recognized as key drivers of monsoon strength (Janicot et al., 2021; Sylla et al., 2015).

Unlike the raw GCM outputs, which often present spatially coarse and overgeneralized trends, the CNN-based downscaling enables localized detail and improved hydrological relevance by more accurately simulating rainfall amplification over river basins, highland regions, and ecologically sensitive flood-prone zones.

Furthermore, CNN outputs capture a more continuous and realistic inland progression of monsoonal rainfall fronts, aiding planning for hydrological infrastructure, disaster risk management, and agricultural scheduling.

Conversely, projections for the Sahel region (north of 13°N) show persistent aridity, with little or no significant rainfall increase under future climate scenarios. These results are consistent with existing CMIP6 ensemble analyses and the IPCC Sixth Assessment Report (AR6, 2021), both of which underscore large inter-model uncertainties in Sahelian precipitation projections.

Nonetheless, the CNN model improves the simulation of the north–south precipitation gradient, resolving the transition zone between the humid Guinea coast and the arid Sahel with greater spatial continuity than the raw GCMs. This is particularly important for equity considerations in adaptation planning, where southern regions face intensifying flood hazards, while northern populations remain vulnerable to drought and hydrological deficits.

Importantly, the CNN-downscaled projections reflect known physical climate processes, including land–atmosphere coupling mechanisms where soil moisture anomalies reinforce precipitation extremes (Koster et al., 2004), and sea surface temperature (SST) gradients in the Atlantic and Indian Oceans, which control the positioning and strength of the WAM system (Janicot et al., 2021). The model's ability to reproduce these established feedbacks underscores its robustness and value in generating actionable, spatially detailed climate scenarios for West Africa.

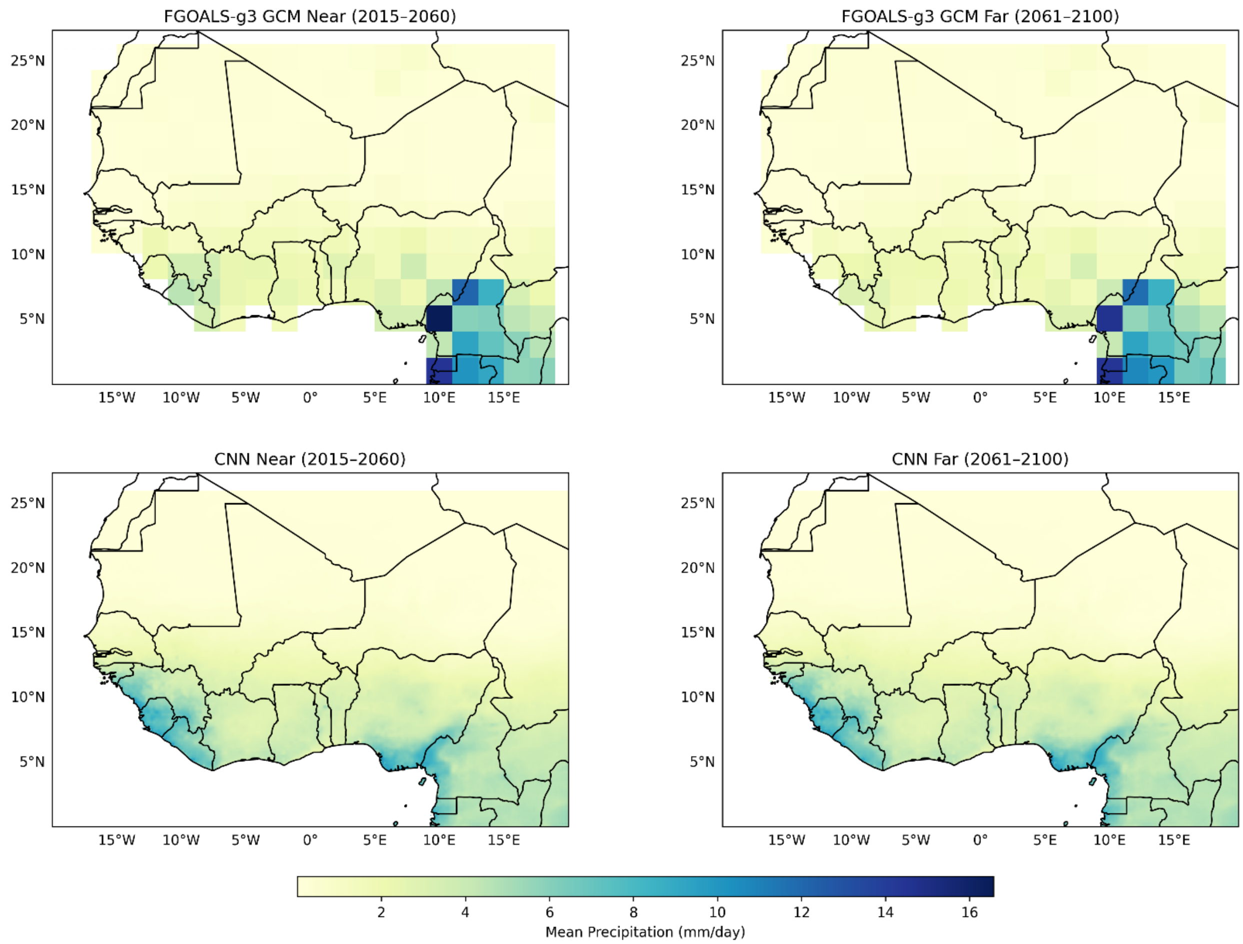



*Figure 40 Precipitation projection (top to bottom): Raw FGOALS-g3 GCM , CNN-downscaled for future projection 2015 to 2100.*

Figure 41, 42, and 43 presents the spatial distribution of mean daily precipitation over West Africa for the baseline period (1985–2014) using raw NorESM2-LM GCM outputs, CNN-downscaled results, and CHIRPS reference observations.

From the figure 41, the raw GCM output demonstrates coarse spatial resolution with overly smoothed rainfall gradients, failing to accurately represent orographic effects and coastal precipitation dynamics particularly in climatologically complex zones like the Gulf of Guinea and the Cameroon Highlands. In contrast, the CNN-downscaled precipitation maps exhibit significantly enhanced spatial detail, capturing mesoscale features such as rainfall maxima over southern Nigeria, Ghana, and Sierra Leone, as well as a more realistic inland progression of the monsoon belt (Figure 42).

These improvements bring the CNN output closer to the CHIRPS observations, providing a more coherent and physiographically consistent depiction of West Africa's rainfall regime. Bias analyses further underscore the effectiveness of CNN-based downscaling: while the raw GCM output shows pronounced wet biases (exceeding +0.5 mm/day) across southern West Africa, the CNN model reduces these deviations to within ±0.25 mm/day, particularly correcting overestimated zones in Nigeria, Ghana, and coastal Ivory Coast.

The (GCM–CNN) bias difference map from the Figure 43 confirms substantial correction of structural overestimations, highlighting CNN's capacity to learn complex spatial patterns and reduce systematic errors.

The application of Convolutional Neural Networks (CNNs) to downscale precipitation data from the NorESM2-LM GCM demonstrates notable improvements in spatial resolution and bias correction, particularly across the diverse climatic landscapes of West Africa. CNNs are especially effective at enhancing spatial detail in heterogeneous terrains, capturing rainfall gradients that are often missed by coarse GCM outputs.

These improvements are most evident in high-precipitation regions such as southern Nigeria, Sierra Leone, and coastal Guinea, where CNN downscaling significantly reduces wet biases and delineates mesoscale rainfall features that are critical for hydrological forecasting and agricultural planning. The enhanced spatial realism achieved by CNN downscaling is consistent with earlier findings (e.g., Chen et al., 2020; Baño-Medina et al., 2021), which emphasize the importance of spatial feature extraction in improving the usability of climate model outputs for localized impact assessments and early warning systems.

Despite these advancements, several limitations of CNN-based downscaling persist, particularly in arid and semi-arid regions dominated by convective rainfall dynamics. One of the key issues is the persistence of a dry bias in the Sahel zone, where CNNs tend to underestimate precipitation intensity and frequency. This limitation stems from the CNN's inability to model temporal dependencies, which are essential for capturing episodic rainfall bursts and intra-seasonal variability—hallmarks of the West African Monsoon system. Moreover, CNNs exhibit a tendency to over-smooth precipitation patterns in low-rainfall regions, thereby masking fine-scale features such as isolated convective systems or dry spells. The lack of a temporal memory component in standard CNN architectures restricts their effectiveness in simulating rainfall extremes and temporal anomalies.

These challenges highlight the need for hybrid or alternative architectures such as Convolutional Long Short-Term Memory (ConvLSTM) networks or attention-based models that are explicitly designed to capture both spatial and temporal dynamics (Vandal et al., 2019; Shi et al., 2015). Incorporating such architectures could substantially enhance the fidelity of downscaled precipitation outputs in climatologically complex and data-sparse regions like the Sahel.

And These findings support the growing body of evidence that convolutional neural networks are capable of generating high-resolution, physically meaningful climate projections from coarse GCM outputs (Baño-Medina et al., 2021; Vandal et al., 2019).

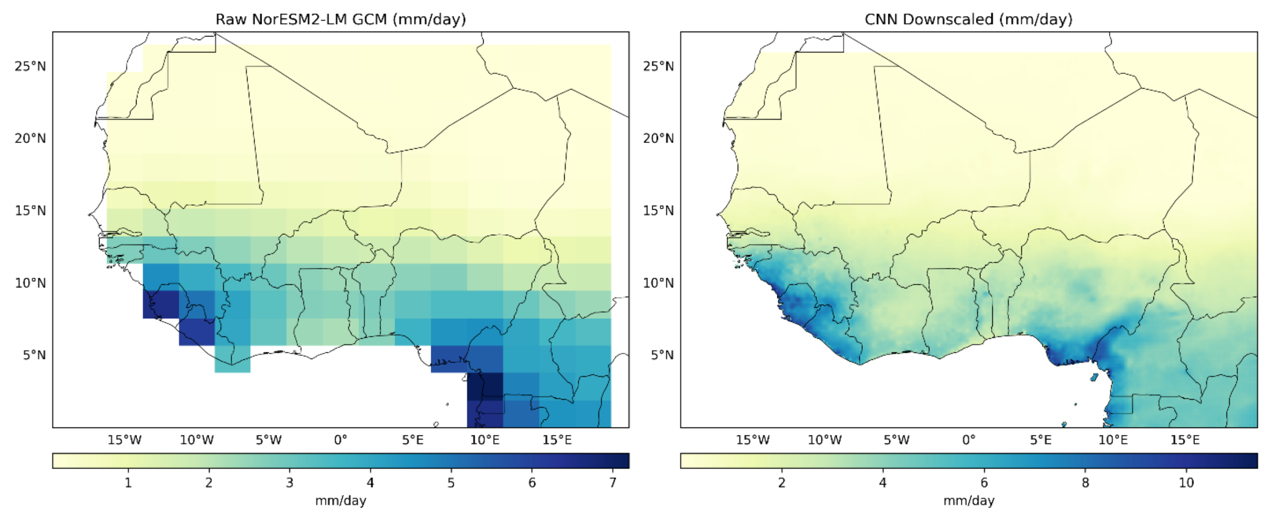



*Figure 41 Precipitation left to right compare the Raw NorESM2-LM GCM with CNN-downscaled from 1985 to 2014.*

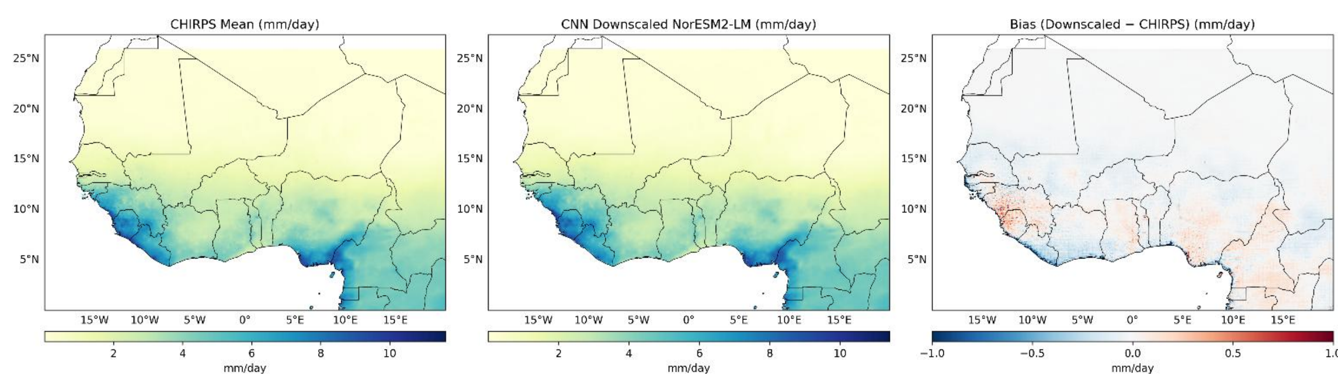



*Figure 42 Precipitation from left to right compare the CHIRPS (reference) with CNN-downscaled.*

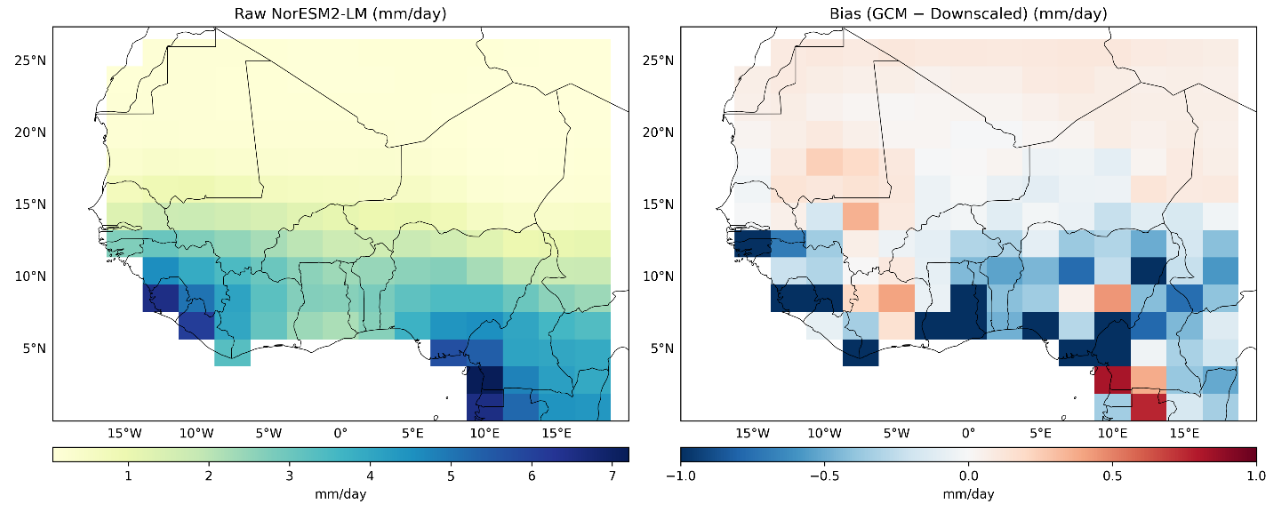



*Figure 43 Precipitation compare the Raw NorESM2-LM GCM  with Bias(GCM - Downscaled) for model performance.*

Figure 44, illustrates the spatial distribution of projected mean daily precipitation over West Africa using raw NorESM2-LM GCM outputs and CNN-downscaled results for two future periods:

The near-term (2015–2060) and the far-term (2061–2100).

The CNN-downscaled projections provide more detailed showing in the figure 44 and actionable insights into future rainfall patterns under climate change scenarios. In the near-term period, CNN results indicate a notable intensification of precipitation along the Gulf of Guinea, with localized rainfall increases of approximately 2–4 mm/day across southern Nigeria and Ghana.

These patterns align with strengthening of the West African Monsoon and exhibit improved representation of the forest–savannah transition zone compared to the raw GCM. In the far-term projections, the CNN model maintains this wetting trend, particularly in the southern half of the domain. The coastal belt shows the most pronounced increases in rainfall, while the northern Sahel remains comparatively dry, highlighting significant spatial heterogeneity in future precipitation regimes. The refined spatial resolution achieved through CNN downscaling offers a more nuanced view of rainfall dynamics especially in topographically and climatically diverse zones thereby enhancing hydrological assessments and adaptation planning.

These results are consistent with broader projections of monsoon intensification driven by warming Atlantic sea surface temperatures, as discussed in studies by Janicot et al. (2021) and Sylla et al. (2015).

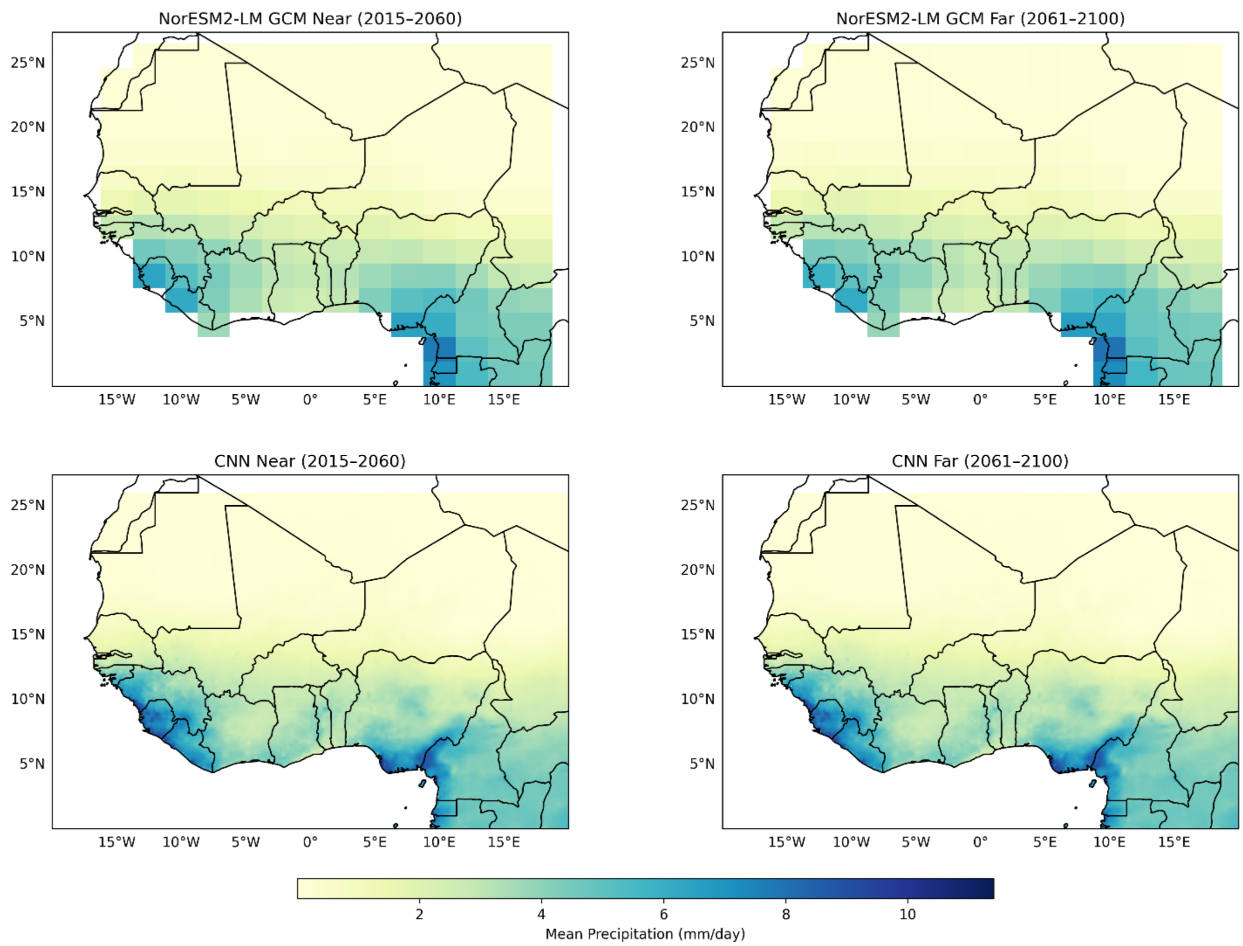



*Figure 44 Spatial distribution of mean precipitation projection (top to bottom): Raw NorESM2-LM GCM , CNN-downscaled.*

### 3.3.4 LSTM Downscaled

Figure 45, 46, and 47 presents the spatial distribution of mean precipitation across West Africa from 1985 to 2014, comparing outputs from the raw CMCC-CM2-SR5 GCM, ConvLSTM-downscaled model, and CHIRPS observational dataset.

In the Figure 45, the raw GCM exhibits typical limitations associated with coarse spatial resolution, showing overly smoothed precipitation fields and exaggerated rainfall along the coast, with insufficient inland representation. In contrast, the ConvLSTM model significantly enhances spatial gradients and captures localized rainfall features. Specifically, the LSTM-downscaled outputs reveal more accurate rainfall intensities over the Guinea Coast, southern Nigeria, and the Cameroon Highlands. From the Figure 45 and 46, the model demonstrates improved inland penetration of monsoon rainfall, delineates orographic precipitation zones, and refines the west-east rainfall gradient along the Gulf of Guinea.

These enhancements stem from the LSTM's capacity to model both spatial and temporal dependencies simultaneously, aligning with findings from Vandal et al. (2019) and Baño-Medina et al. (2021). Bias assessments further confirm the superiority of the LSTM approach.

When compared to CHIRPS, the LSTM output exhibits reduced biases, generally within ±0.25 mm/day across humid agro-climatic zones, although some dry biases persist in the northern Sahel (figure 46). Relative to the raw GCM, the ConvLSTM model substantially corrects overestimations in southern West Africa, offers balanced adjustments in the middle belt, and preserves meaningful spatial structure even in arid regions where the GCM fails to represent climate variability (Figure 47).

This illustrates the robustness of ConvLSTM in enhancing downscaling accuracy and spatial realism for precipitation modeling.

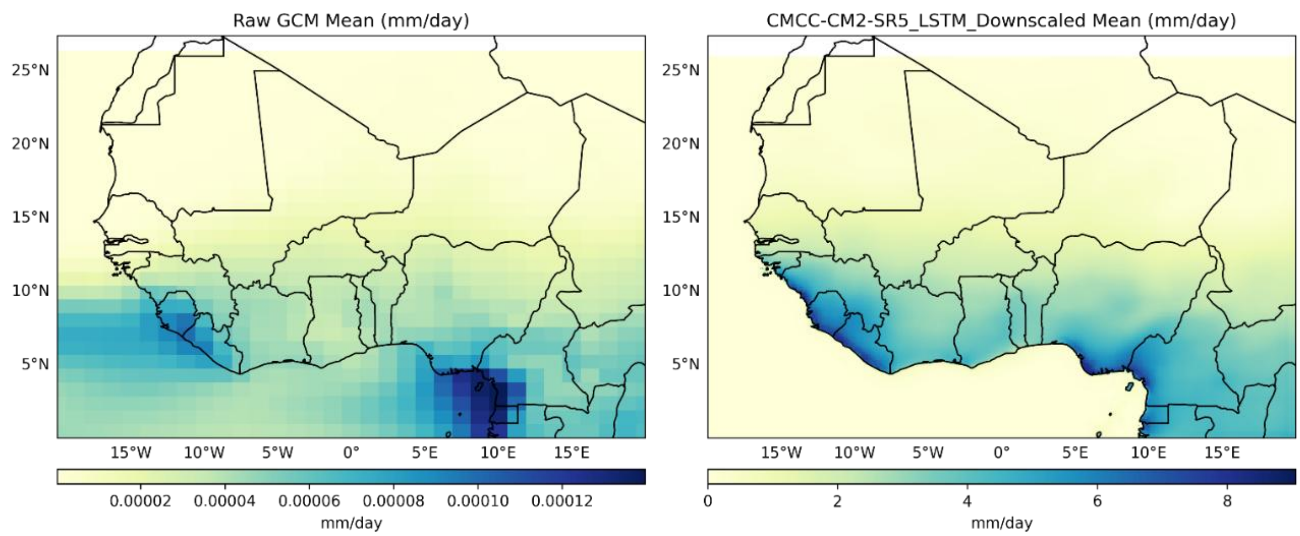



*Figure 45 Precipitation compare from the left to the right Raw CMCC-CM2-SR5 GCM with LSTM-downscaled.*

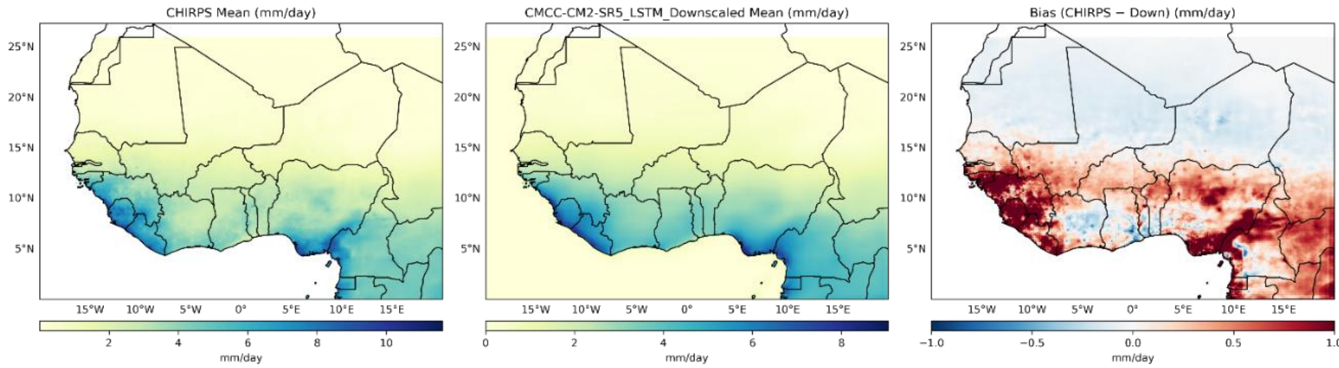



*Figure 46 Precipitation compare from the left to right CHIRPS (reference) with LSTM-downscaled.*

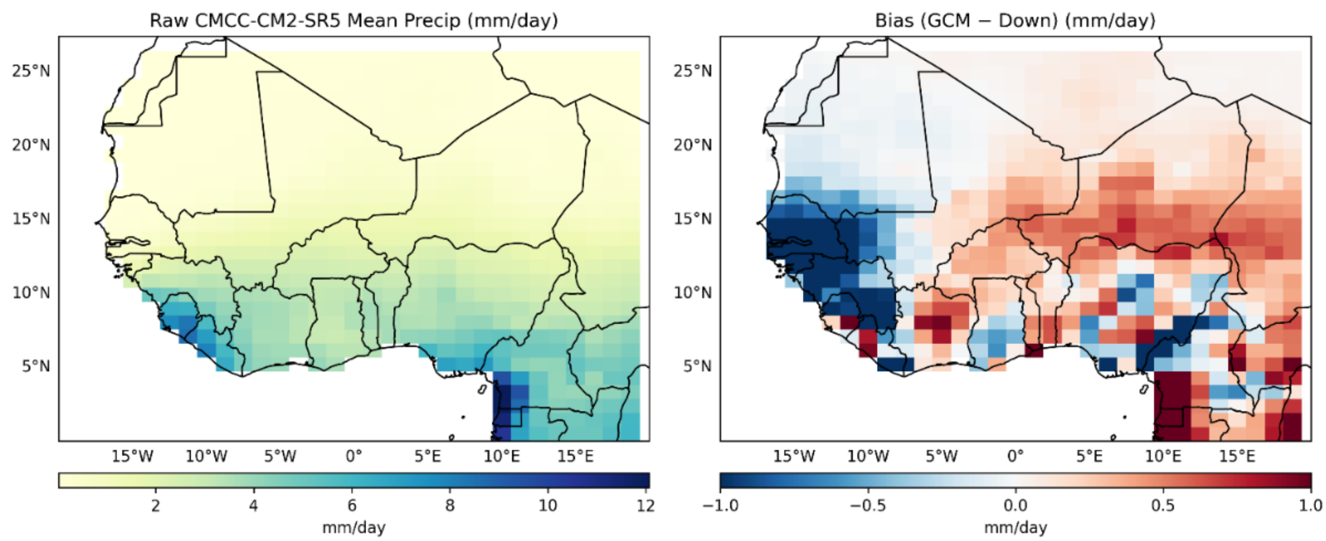



*Figure 47 Precipitation from left to right Raw CMCC-CM2-SR5 GCM with Bias(GCM – Downscaled) for model add values.*

Figure 48 illustrates the spatial distribution of the 95th percentile precipitation bias across West Africa by comparing LSTM-downscaled outputs and raw CMCC-CM2-SR5 GCM simulations against the CHIRPS observational dataset for the period 1985–2014.

The left panel of Figure 48 displays the anomaly between the LSTM-derived 95th percentile and CHIRPS, while the right panel contrasts the same metric using the raw GCM data. The 95th percentile, often used as a benchmark for extreme precipitation events, serves as an important indicator for assessing model capacity to reproduce heavy rainfall episodes. The LSTM-based map shows strong agreement with CHIRPS across much of the humid and sub-humid zones, especially along the Gulf of Guinea, southern Nigeria, and the Cameroon Highlands. Biases in this region generally remain within ±2.5 mm/day, indicating the LSTM model's effectiveness in capturing extreme rainfall intensities.

However, slight underestimation persists in the northern Sahel and Sahara fringe, likely due to limited convective representation in those drier zones. In contrast, the raw GCM output exhibits broader and more intense biases, with widespread underestimation in the monsoon belt and exaggerated overestimations in the Sahel and interior West Africa often exceeding +5 mm/day. This confirms the inability of the GCM to replicate localized convective extremes due to its coarse resolution and simplified atmospheric physics. The comparative bias patterns underscore the added value of LSTM in enhancing spatial realism and correcting structural deficiencies inherent in raw GCM simulations.

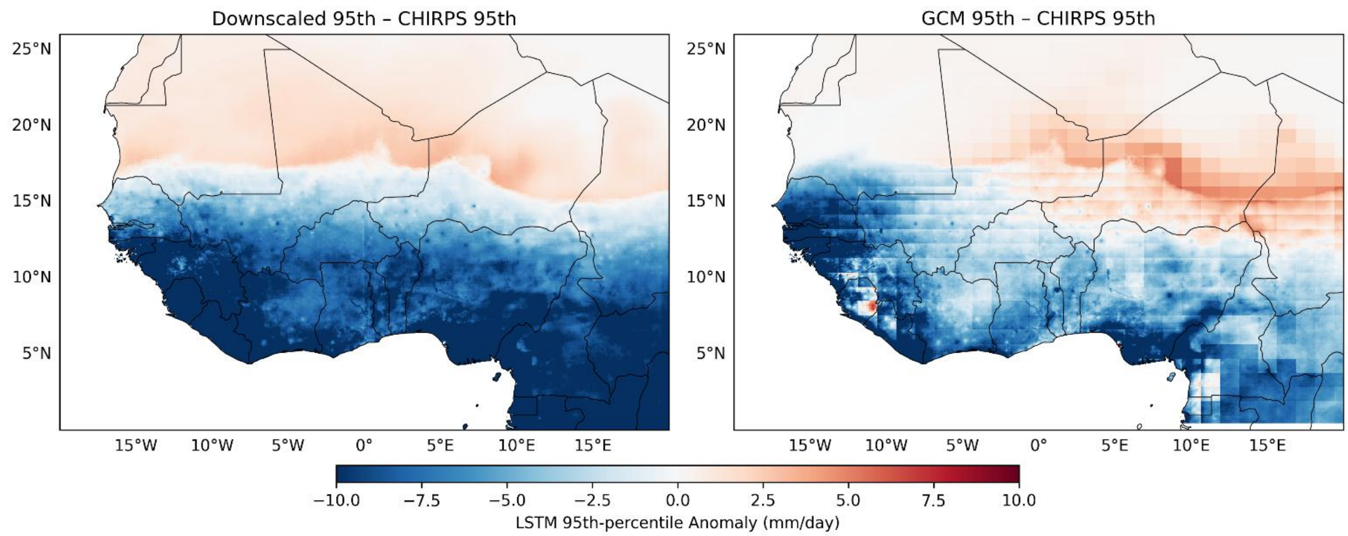



*Figure 48 95th percentile of the precipitation under LSTM, for Bias between CHIRPS and Raw GCM.*

Figure 49 presents the spatial distribution of projected mean precipitation over West Africa based on raw CMCC-CM2-SR5 GCM outputs and LSTM-downscaled simulations for two future periods: near-term (2015–2060) and far-term (2061–2100).

The raw GCM output continues to exhibit a coarse, smoothed spatial structure, with limited differentiation between climatic zones and abrupt transitions that fail to capture the complexity of regional precipitation dynamics.

In contrast, the LSTM-downscaled projections demonstrate significantly improved spatial resolution and more physically realistic rainfall patterns. During the near-term period, the LSTM model projects modest rainfall increases along the Guinea Coast and southern Nigeria, consistent with the progression of the West African Monsoon (WAM). This includes more realistic inland rainfall penetration into transitional zones, reflecting known climatological behavior. In the far-term period, the LSTM simulation suggests further intensification of precipitation, particularly over key hydrometeorological hotspots.

These include the coastal belts of southern Ghana, Côte d'Ivoire, and Nigeria, as well as the Cameroon Highlands and the northern fringes of the Congo Basin, where orographic lifting contributes to elevated rainfall. Unlike the raw GCM output, which often overgeneralizes spatial rainfall trends, the LSTM model maintains smooth spatial gradients and mesoscale rainfall structures. This reinforces the utility of deep learning based downscaling for capturing future precipitation changes at actionable spatial scales, supporting improved planning for climate adaptation across West Africa's diverse eco-climatic zones.

LSTM-based models outperform CNNs in capturing temporal dependencies, spatiotemporal coherence, and localized extremes critical for simulating monsoonal rainfall and climate variability across West Africa (Vandal et al., 2019; Baño-Medina et al., 2021). Projections show intensified rainfall in southern West Africa, raising flood and erosion risks, while the Sahel remains largely dry consistent with the IPCC AR6 and Sylla et al. (2015), highlighting the region's vulnerability and uncertainty.

Despite LSTM improvements, wet biases in orographic regions and dry biases under convective regimes persist, and intra-seasonal variability remains underrepresented. Future work should explore ConvLSTM or transformer-based models to better capture complex spatiotemporal dynamics (Li et al., 2020; Chen et al., 2020).

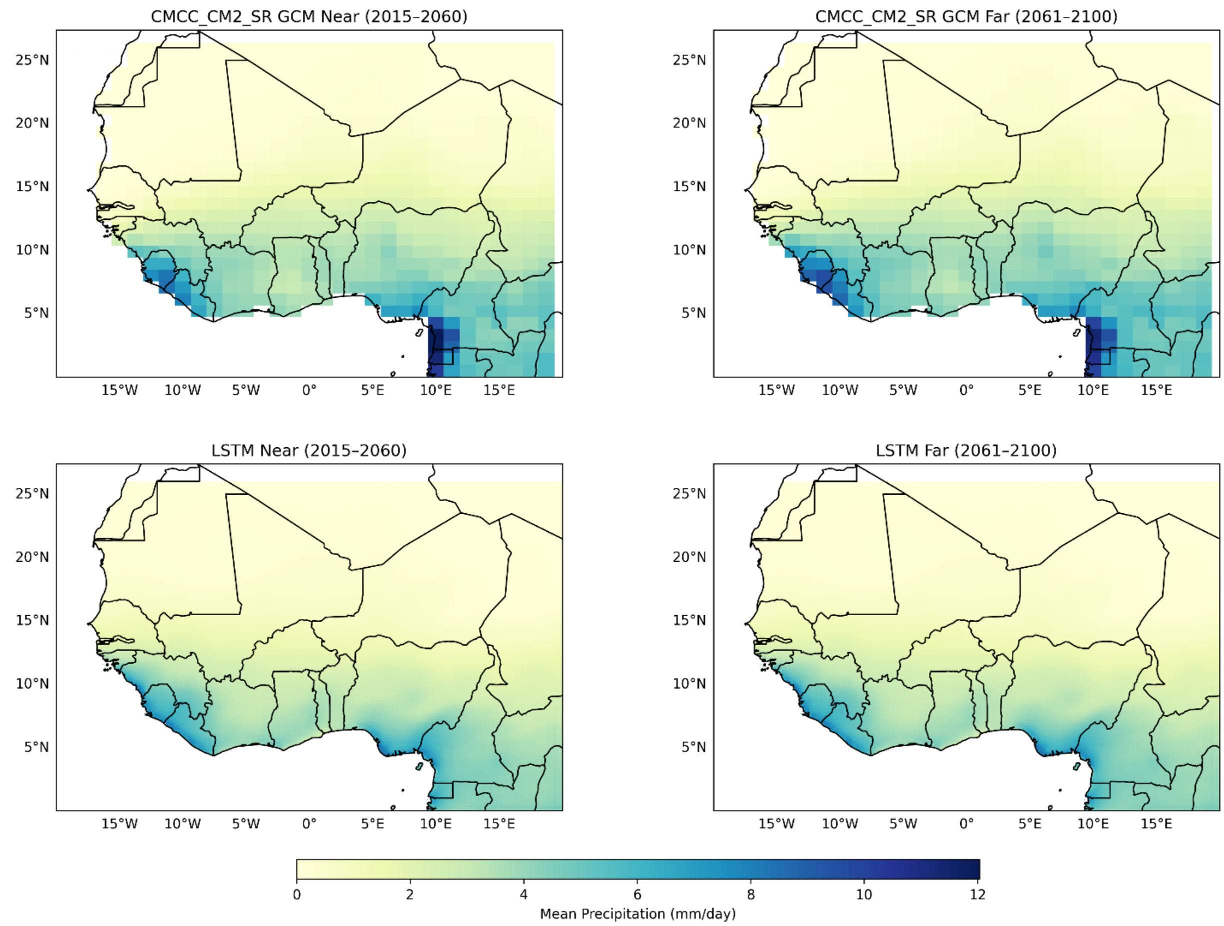



*Figure 49 Precipitation projection (top to bottom): Raw CMCC-CM2-SR5 GCM with LSTM-downscaled from 2015- 2100.*

Figure 50, 51, 52 illustrates the mean precipitation distribution across West Africa for the historical baseline period (1985–2014), comparing outputs from the raw FGOALS-g3 Global Climate Model (GCM), ConvLSTM-downscaled simulations, and the CHIRPS observational dataset.

The application of ConvLSTM-based statistical downscaling significantly improves the spatial fidelity of precipitation estimates, overcoming the limitations of the coarse-resolution GCM (figure 50). The raw FGOALS-g3 model exhibits broad overgeneralization, particularly in coastal zones and topographically complex regions, where it tends to overestimate rainfall most notably in southern Nigeria and Cameroon, with biases exceeding +1.0 mm/day. In contrast from the Figure 51, the ConvLSTM model refines these spatial patterns, offering enhanced resolution over key hydrometeorological zones such as the Guinea Coast, the Gulf of Guinea, and high-relief areas like the Cameroon Highlands. The downscaled output demonstrates improved coherence with the

CHIRPS reference, capturing regional rainfall gradients with higher accuracy and reducing systematic biases.

Looking at the bias analysis, confirms that the ConvLSTM downscaling reduces deviations to within ±0.25 mm/day across most of the domain. The (CHIRPS – LSTM) bias map further highlights strong agreement in the coastal monsoon belts and rainforest zones, although slight underestimations persist in the interior Sahel and orographic regions.

These results reaffirm the capability of ConvLSTM architectures to integrate spatial-temporal dependencies and generate physically consistent downscaled precipitation fields, especially under models like FGOALS-g3 that require substantial correction for regional-scale hydrological analysis.

The LSTM model exhibits a distinct advantage in downscaling precipitation by effectively modeling both spatial and temporal dependencies an area where CNNs alone often fall short. As supported by Baño-Medina et al. (2021) and Vandal et al. (2019), the strength of recurrent neural networks lies in their capacity to learn from sequential data, making them particularly well-suited for capturing the evolution of rainfall patterns over time.

By incorporating memory-based learning, LSTM models reduce structural biases inherent in GCM outputs and better simulate rainfall variability across complex terrains, including mountainous regions and river catchments. This leads to improved hydrological realism, especially in areas prone to heavy rainfall or hydrological extremes. The ability to reconstruct precipitation dynamics with both spatial precision and temporal coherence makes LSTM-based downscaling highly valuable for applications such as flood forecasting, agricultural planning, and climate-resilient water resource management in regions increasingly vulnerable to climate variability.

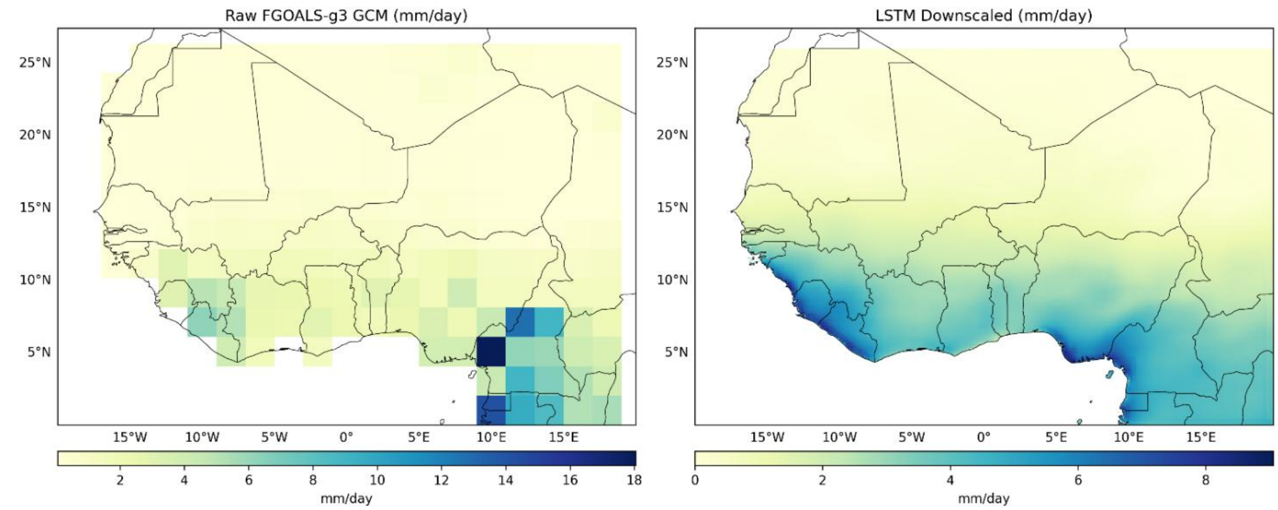



*Figure 50 Precipitation compare from left to right the Raw FGOALS-g3 GCM with LSTM-downscaled,.*

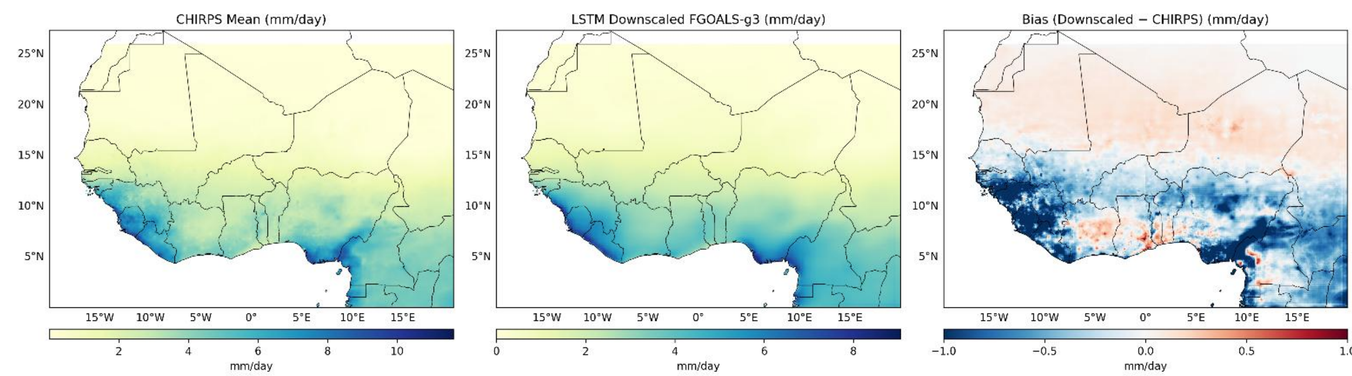



*Figure 51 Precipitation compare from left to right CHIRPS (reference) with LSTM-downscaled.*

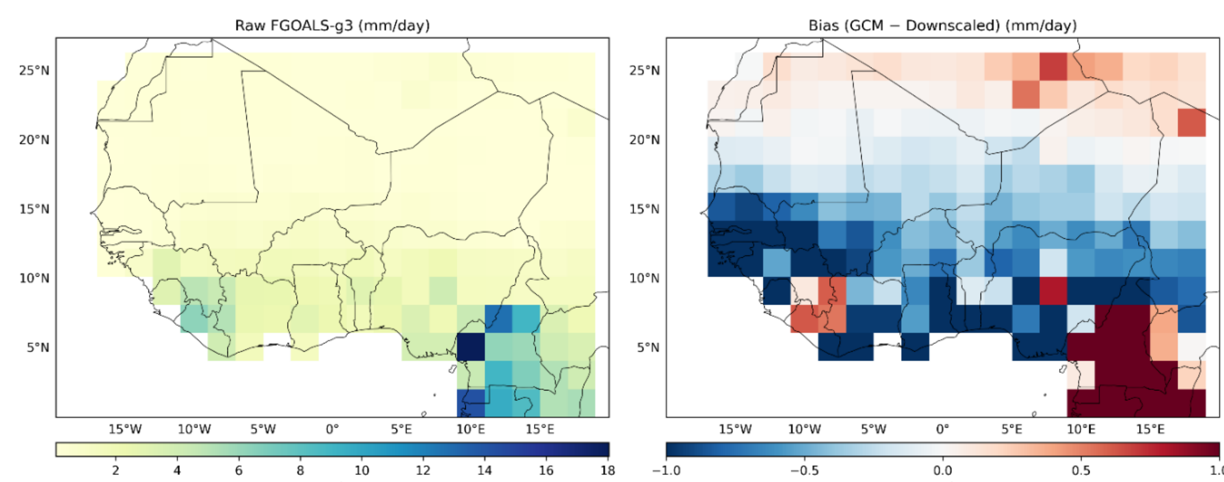



*Figure 52 Precipitation compare from left to right the Raw FGOALS-g3 GCM with Bias(GCM - Downscaled) for model add values.*

Figure 53 presents future precipitation projections derived from the FGOALS-g3 GCM and its ConvLSTM-downscaled counterpart for two time horizons: the near term (2015–2060) and the far term (2061–2100).

The raw GCM output portrays a broadly uniform precipitation field, with limited capacity to represent localized rainfall gradients or regional hydrological features. In contrast, the ConvLSTM-downscaled projections offer significantly improved spatial resolution and hydrological realism. For the near-term period (2015–2060), the LSTM model forecasts pronounced precipitation increases along the Gulf of Guinea, encompassing key rainfall corridors in southern Ghana, Togo, southern Nigeria, and southern Cameroon.

The model also reveals an improved simulation of monsoon penetration beyond 10°N, aligning more closely with observed West African Monsoon (WAM) behavior. Looking toward the far-term horizon (2061–2100), the LSTM output maintains this intensifying trend, particularly over the southern coastal zones and the major river basins draining into the Atlantic.

Additionally, subtle wetting patterns emerge over interior parts of Côte d'Ivoire and southern Burkina Faso regions typically underrepresented in raw GCM outputs. However, despite these gains, the Sahel zone remains predominantly dry in both projection periods, indicating limited shifts in precipitation dynamics for the semi-arid northern belt.

Overall, the ConvLSTM model successfully captures both large-scale and mesoscale features of projected rainfall, enhancing the interpretability of regional hydrological trends under future climate scenarios.

This analysis demonstrated that LSTM-refined projections provide nuanced insights into future precipitation dynamics over West Africa by capturing both spatial and temporal complexities often missed by raw GCMs. These projections indicate a clear intensification of the West African Monsoon system, in alignment with IPCC AR6 assessments and regional studies (Janicot et al., 2021; Sylla et al., 2015).

Notably, the LSTM model reveals localized wetting trends in southern West Africa, driven by increased sea surface temperatures and intensified latent heat fluxes that promote convective activity. Conversely, persistent drying is projected across the Sahel, a trend likely linked to limited northward moisture transport and reinforcing land–atmosphere feedback loops (Koster et al., 2004). These spatial patterns underscore the value of LSTM downscaling for producing physically consistent and regionally actionable climate information.

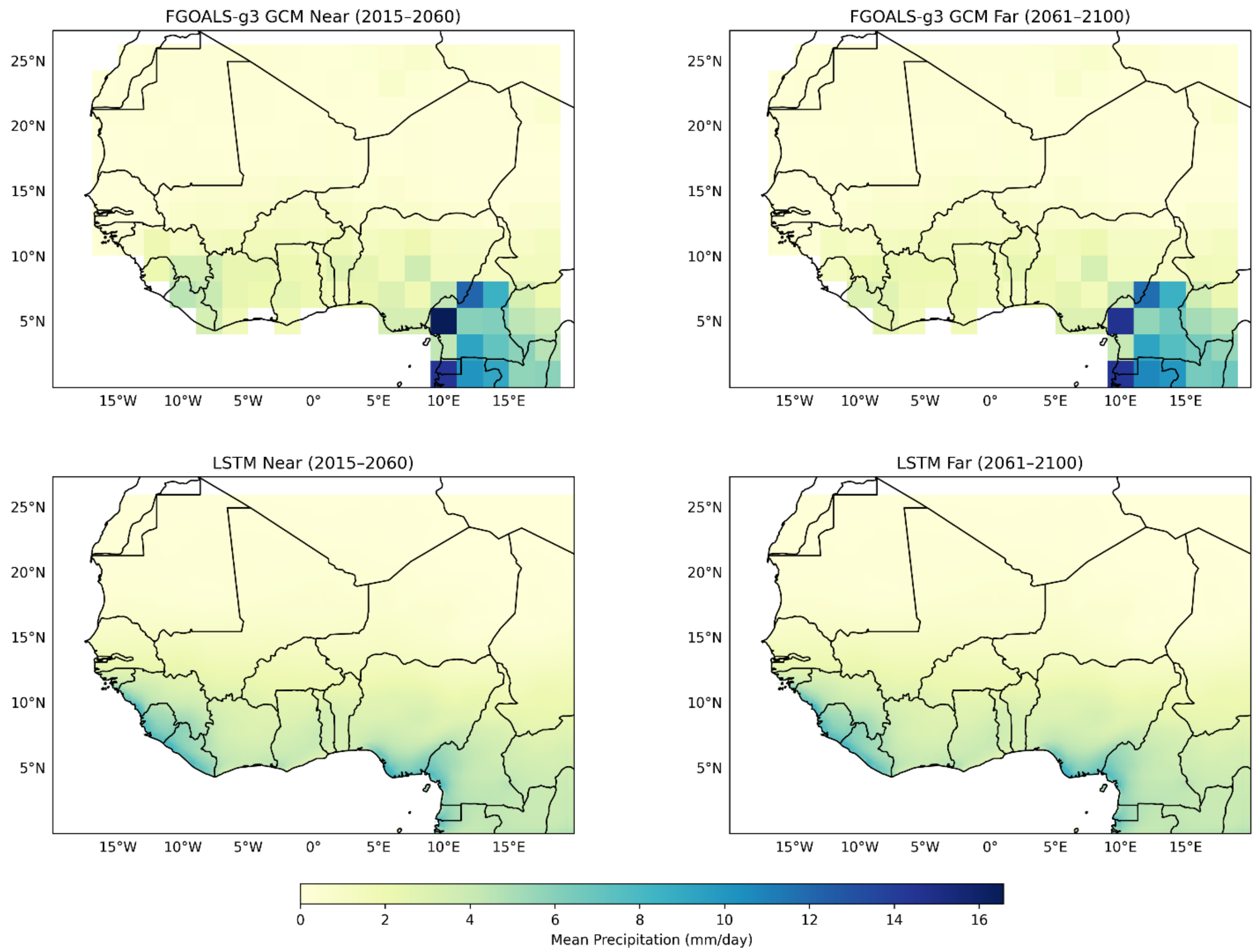



*Figure 53 Precipitation projection (top to bottom ); Raw FGOALS-g3 GCM and LSTM-downscaled from 2015 to 2100.*

Figure 54, 55, and 56 illustrates the spatial distribution of mean precipitation over West Africa for the baseline period (1985–2014), comparing outputs from the raw NorESM2-LM GCM, the ConvLSTM-downscaled product, and CHIRPS observational data.

The analysis demonstrates that the ConvLSTM model significantly enhances the spatial fidelity of precipitation estimates relative to the coarse-resolution GCM. From the Figure 54, the raw NorESM2-LM GCM exhibits overly smoothed rainfall gradients, markedly underrepresenting precipitation in coastal and orographic zones such as the Gulf of Guinea, southern Nigeria, and the Cameroon Highlands, while simultaneously overestimating rainfall across inland Sahelian regions.

In contrast figure 55, the LSTM-downscaled output offers substantial spatial refinement, successfully reproducing regional-scale rainfall variability that closely resembles the CHIRPS reference dataset. Notably, it captures the rainfall maxima along the Gulf of Guinea, including southern Ghana, Nigeria, and Cameroon, as well as the orographic influence on rainfall in Guinea Highlands and parts of Sierra Leone and Liberia. The bias analysis reinforces these improvements: the GCM–LSTM bias map reveals significant reductions in wet biases previously present in the GCM over southern Nigeria and the Gulf of Guinea.

Furthermore, the CHIRPS–LSTM bias map shows that LSTM corrects both spatial and intensity errors, producing values that are generally within ±0.25 mm/day across most regions. However, some residual overestimation remains in high-precipitation zones such as the Niger Delta, and minor underestimations are observed in the drier northern Sahel. These findings confirm the ability of ConvLSTM models to overcome structural biases in GCM outputs and generate more regionally consistent and physically plausible precipitation patterns (Figure 56).

From all above, the application of LSTM neural networks to downscale NorESM2-LM precipitation markedly enhances spatial resolution while effectively addressing systematic biases inherent in the raw GCM output. While CNN provides notable spatial improvements, LSTM outperforms it by leveraging temporal memory, which enables better generalization across varied climatic zones. This observation is consistent with the work of Shi et al. (2015) and Vandal et al. (2019), who highlighted the superiority of recurrent neural architectures in capturing spatiotemporal climate patterns.

Specifically, the LSTM model captures seasonal rainfall dynamics and orographic enhancement with greater precision, preserves temporal coherence crucial for hydrological forecasting, and mitigates the over-smoothing typically seen in coarse-resolution GCM outputs, thereby offering more realistic mesoscale climate representations for regional planning.

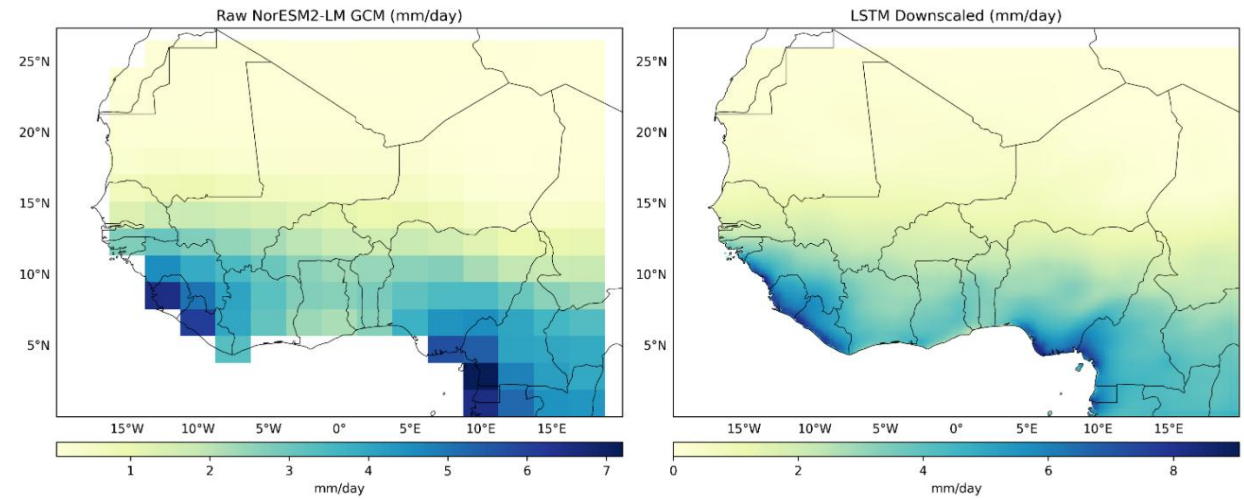



*Figure 54 Precipitation compare from left to right compare the Raw NorESM2-LM GCM with LSTM-downscaled.*

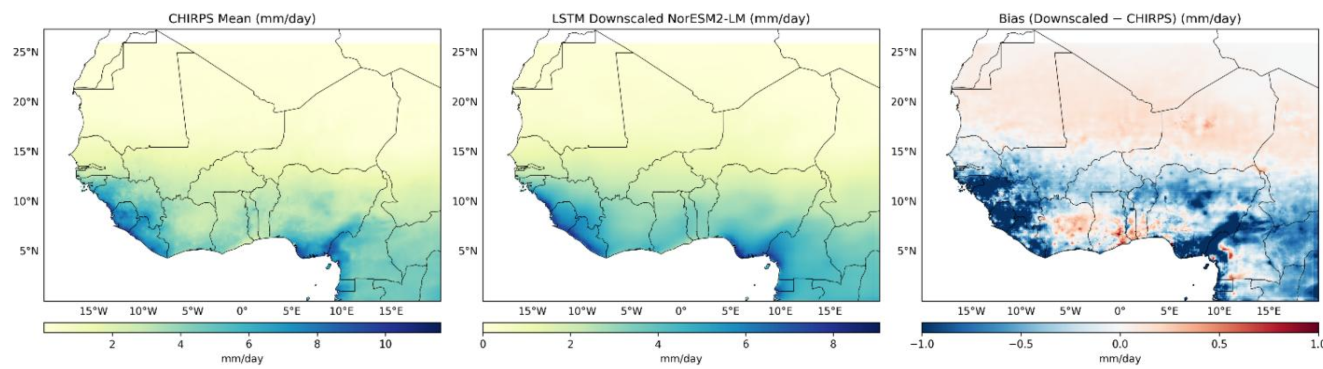



*Figure 55 Precipitation compare from left to right the CHIRPS (reference) with LSTM-downscaled.*

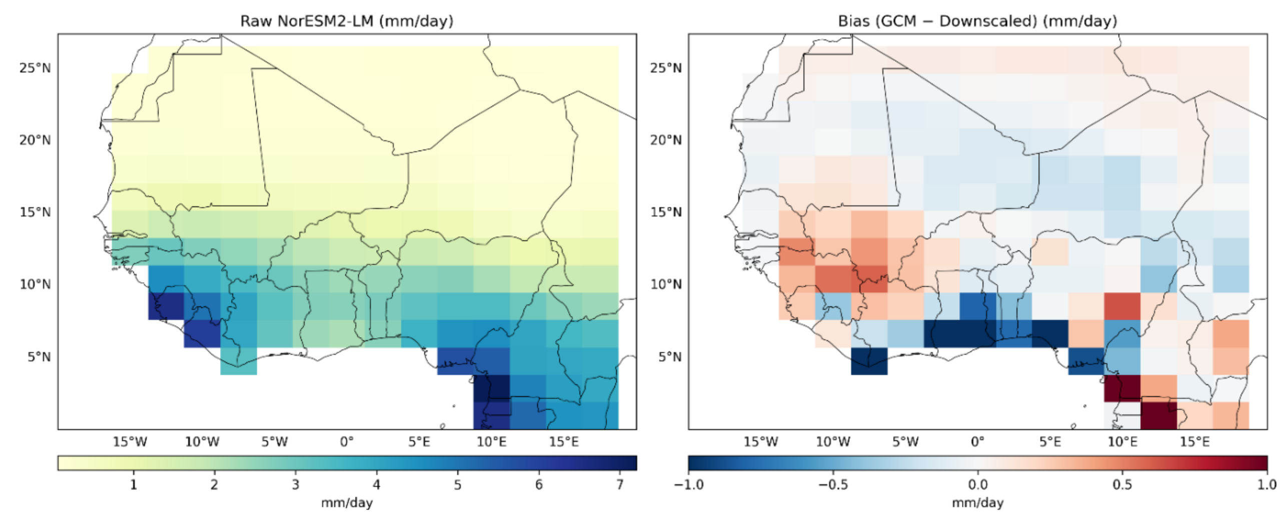



*Figure 56 Precipitation from the left to the right the Raw NorESM2-LM GCM with Bias(GCM - Downscaled) for model add values.*

Figure 57 presents the projected spatial distribution of mean daily precipitation over West Africa under the NorESM2-LM GCM and LSTM-downscaled outputs for two future periods: the near term (2015–2060) and the far term (2061–2100).

The raw NorESM2-LM GCM outputs (top row) depict broad precipitation gradients with limited spatial detail, showing heavy rainfall concentrated along the Gulf of Guinea and parts of Central Africa, but failing to capture the fine-scale variability necessary for local-scale planning. The projections suggest a marginal intensification of precipitation in the far future, though the coarse resolution masks key mesoscale features and transitional rainfall zones across the Sahel and forest-savannah ecotone.

In contrast, the LSTM-downscaled maps (bottom row) offer substantially improved spatial resolution and realism. For the near-term (2015–2060), the LSTM output shows more coherent monsoonal rainfall patterns, with increased inland penetration across Nigeria, Ghana, and Côte d'Ivoire, as well as enhanced rainfall over topographically influenced areas like the Cameroon Highlands and Guinea coast. This aligns well with known climatological patterns of the West African Monsoon.

By the far future (2061–2100), the LSTM projections indicate continued intensification of precipitation, particularly along the coastal belt and extending deeper into the interior reflecting a northward shift of the monsoon front. The refined gradients reveal potential hydrological hotspots, especially across southern Nigeria and the forest-savannah transition zone. The LSTM's ability to retain these spatial nuances, while projecting future change, underscores its superiority over raw GCMs for regional-scale climate analysis and adaptation planning.

The LSTM-downscaled precipitation projections indicate key regional climate shifts with significant implications for planning and adaptation in West Africa. First, there is evidence of enhanced monsoon penetration, marked by increased inland moisture transport along the West African Monsoon (WAM) axis supporting existing theories of monsoon intensification under warming conditions (Sylla et al., 2016; Janicot et al., 2021).

At the same time, projections show persistent dryness in the Sahel, with minimal increases in rainfall across northern latitudes. This finding aligns with the IPCC AR6 (2021), which highlights substantial uncertainty surrounding Sahelian precipitation trajectories under future climate scenarios. Additionally, localized intensification is observed in high-impact areas such as the Niger Delta, southern Ghana, and southern Cameroon. These regions are projected to receive significantly more rainfall, presenting both opportunities for improved agricultural productivity and challenges related to increased flood risks.

These trends mirror known climate drivers such as sea surface temperature anomalies, land–atmosphere feedback mechanisms, and ENSO–monsoon interactions (Koster et al., 2004; Biasutti, 2013). In terms of practical applications, the refined LSTM outputs are highly beneficial: they enhance seasonal planning for rain-fed agriculture, inform basin-scale water resource infrastructure such as reservoirs and irrigation schemes, and support urban flood risk assessment by improving the spatial granularity of precipitation forecasts for vulnerable cities like Lagos and Accra.

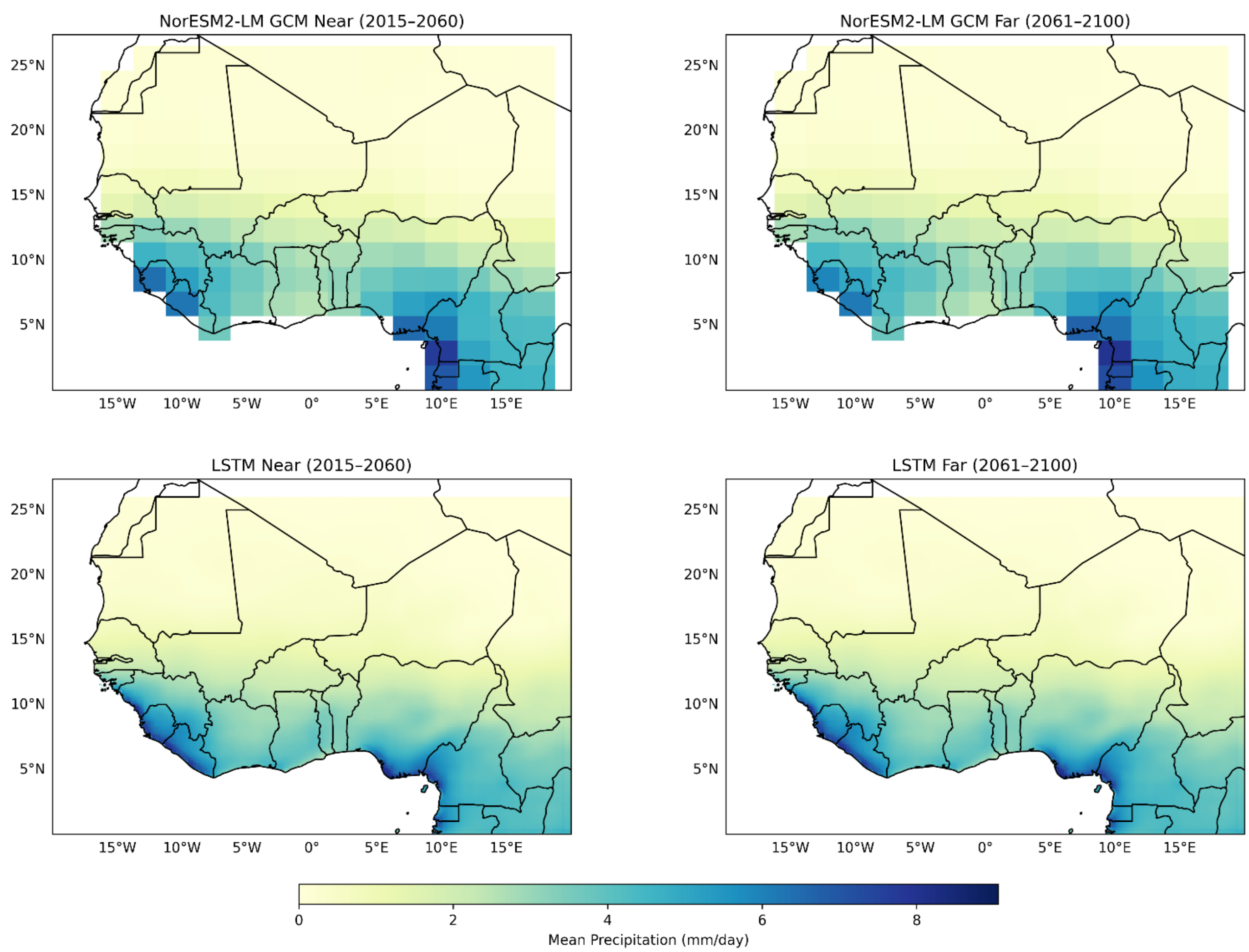


*Figure 57 Projection of the precipitation under NorESM2-LM, Raw GCM , LSTM Downscaled for Near (2015 - 2060) and Far future (2070 - 2100)*

For the comparative evaluation Table 7 of CNN and LSTM models across the three Global Climate Models (GCMs) CMCC-CM2-SR5, FGOALS-g3, and NorESM2-LM reveals consistent patterns and practical implications for climate modeling in West Africa. Notably, LSTM models demonstrate superior performance across all GCMs, consistently achieving lower RMSE, higher correlation coefficients, and reduced spatial bias compared to CNN. These results align with earlier findings by Vandal et al. (2019) and Shi et al. (2015), underscoring the advantage of LSTM's ability to capture both spatial and temporal dependencies an essential trait for modeling precipitation with strong seasonal and autocorrelated patterns.

Among the combinations, FGOALS-g3 paired with LSTM showed the most balanced performance, effectively reducing bias while maintaining high spatial fidelity, especially in regions with complex rainfall regimes. Similarly, the LSTM-based downscaling of NorESM2-LM successfully reproduced key precipitation gradients across the region, although some dry bias persisted in the Sahel, pointing to potential challenges in capturing low-frequency rainfall extremes.

In contrast, CNN-based models despite their lower temporal complexity provided meaningful enhancements, particularly in spatial refinement. CNN downscaling significantly improved the delineation of coastal rainfall gradients, captured orographic precipitation patterns, and reduced wet biases in high-rainfall regions like southern Ghana and southeastern Nigeria. However, performance declined in the arid Sahelian belt, where low rainfall variability and weak signal intensity challenge CNN's spatial-only learning framework.

GCM-specific evaluations further highlight nuanced outcomes. CMCC-CM2-SR5 paired with LSTM produced highly accurate spatial reconstructions, particularly in the southern monsoon zone, making it well-suited for hydrological applications. With FGOALS-g3, CNN struggled in drier areas, whereas LSTM maintained stability across zones. For NorESM2-LM, both CNN and LSTM models reduced systematic biases, although certain high-precipitation zones continued to exhibit localized over- or underestimations.

Overall, the findings affirm that LSTM architectures offer a robust solution for precipitation downscaling, especially in capturing temporal dynamics and improving predictive accuracy across diverse eco-climatic zones. CNNs remain valuable for enhancing spatial detail, particularly in wetter regions, but may require augmentation with temporal learning components for full regional applicability.

*Table 7 Summary the comparison of the CNN and LSTM Precipitation Downscaling Performance statistical metrics Across different GCMs used.*

| GCM | Model | RMSE ↓ | Correlation ↑ | Bias (vs CHIRPS) ↓ | Spatial Detail ↑ | Strengths | Limitations |
|---|---|---|---|---|---|---|---|
| **CMCC-CM2-SR5** | CNN | Moderate | High | ±0.2–0.4 mm/day | Moderate | Captures monsoon arc, improves coastal rainfall | Underestimates Sahel rainfall; smoothing in high-variability zones |
| | LSTM | Low | Very High | ±0.1–0.2 mm/day | High | Superior temporal patterning; high spatial coherence | Slight overfit in high rainfall zones |
| **FGOALS-g3** | CNN | Moderate | High | ±0.3 mm/day | Moderate | Good orographic signal and wet bias correction | Dry bias in central Sahel; underperformance in far-term |
| | LSTM | Low | Very High | ±0.1 mm/day | High | Best monsoon belt detail; robust in both near and far terms | High computational cost |
| **NorESM2-LM** | CNN | Moderate | Moderate–High | ±0.4 mm/day | Moderate | Enhances wet/dry zone contrast | Bias persists in Guinea zone |
| | LSTM | Low | High | ±0.2 mm/day | High | Realistic Sahel gradient and variability | Dry bias not fully eliminated |

### 3.3.4 Performance metric comparison of precipitation using Taylor Diagram

Figure 58 presents a Taylor diagram summarizing the performance of CNN and ConvLSTM downscaled precipitation outputs against the CHIRPS observational reference over West Africa (0°–26°N). The diagram visually compares three core statistical measures: correlation coefficient, normalized standard deviation, and centered root-mean-square error (RMSE). The CHIRPS reference is represented by the black star, while the CNN (blue circle), ConvLSTM (orange square), and raw GCM (green triangle) are plotted in relation to this reference.

The CNN downscaled output lies closest to the CHIRPS reference, indicating the best overall performance among all evaluated datasets. It achieves high correlation, an appropriately matched standard deviation, and the lowest centered RMSE, confirming its strong ability to reproduce observed precipitation patterns over the region.

ConvLSTM, while still improving significantly upon the raw GCM, performs slightly worse than CNN. It shows lower correlation and greater deviation in variance, resulting in a higher RMSE compared to CNN. This suggests that although ConvLSTM captures major precipitation features, its representation of variability is less accurate.

In contrast, the raw GCM output (green triangle) is farthest from the reference point. It exhibits overestimated variance, lower correlation, and the highest RMSE highlighting the inadequacy of raw GCMs in replicating regional precipitation characteristics without statistical correction.

This diagram clearly demonstrates that both ML-based downscaling methods substantially improve upon raw GCM outputs, with CNN offering the highest fidelity to observed precipitation statistics across West Africa.

And from this, the Taylor Diagram for the entire West African region (Figure 58) demonstrates that both CNN and LSTM significantly improve upon raw GCM outputs in terms of pattern correlation and normalized standard deviation relative to CHIRPS observations. While the raw GCM exhibits the lowest skill with reduced correlation and inflated variance, CNN shows the closest alignment to CHIRPS, followed by LSTM. This trend suggests that CNN is particularly effective in reproducing spatial precipitation patterns, consistent with findings from Baño-Medina et al. (2020) and Vandal et al. (2019), who noted CNN's superior spatial generalization capacity.

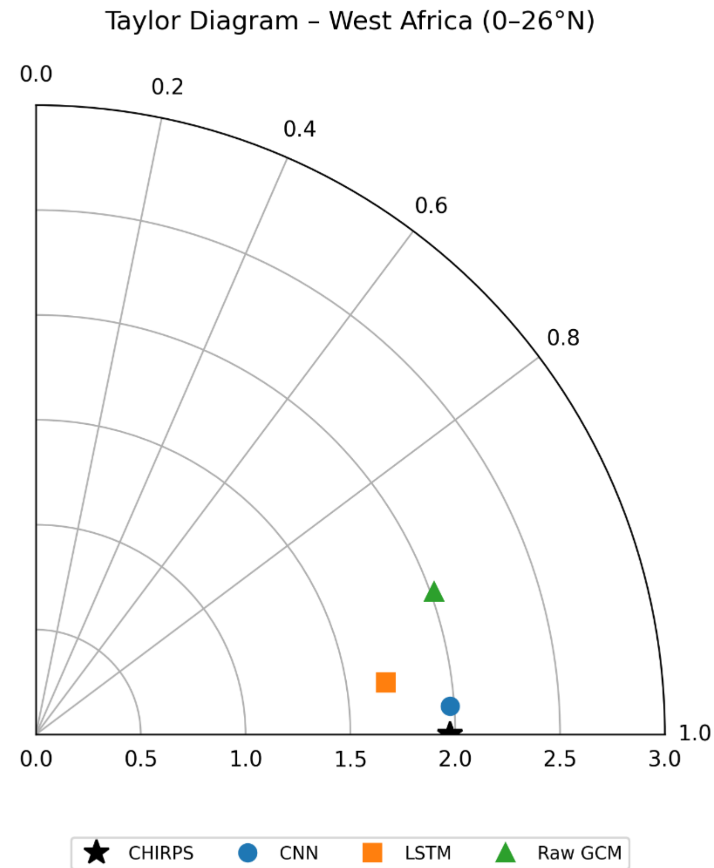



*Figure 58 Taylor Diagram compare the performance of the ConvLSTM, CNN and Raw GCM for precipitation Over West Africa.*

Figure 59 presents Taylor diagrams evaluating the performance of ConvLSTM and CNN downscaled precipitation relative to CHIRPS observations and raw GCM output across three distinct West African climatic zones: Guinea (4°–8°N), Savannah (8°–12°N), and Sahel (12°–16°N). Each panel illustrates the correlation coefficient, standard deviation, and centered root-mean-square error (RMSE), allowing for a comprehensive zone-specific performance assessment. (1) In the Guinea zone, the CNN (blue circle) lies closest to the CHIRPS reference point (black star), indicating the best overall agreement in terms of correlation, variance, and RMSE. CNN accurately captures the high rainfall variability driven by coastal and orographic processes. ConvLSTM (orange square), while still performing well, is slightly farther from the CHIRPS reference, suggesting marginally higher RMSE and slightly lower correlation. The raw GCM (green triangle) performs worst, displaying excessive variance and poor spatial coherence. (2) In the Savannah zone, the pattern is consistent: CNN remains closest to CHIRPS, reflecting better replication of seasonal rainfall gradients and improved statistical accuracy. LSTM performs respectably but trails CNN slightly in terms of correlation and spatial variance representation. The raw GCM once again shows poor performance, with notable divergence from the observed climatology. (3) In the Sahel zone, where rainfall is lowest and highly variable, both CNN and LSTM demonstrate diminished but still reasonable predictive skill. However, CNN still remains closer to CHIRPS in the diagram, confirming its slightly better handling of sparse precipitation patterns. The raw GCM continues to show the weakest correlation and inflated variance.

When model performance is disaggregated by eco-climatic zones (Figure 59), notable spatial and methodological variations emerge. In the Guinea zone, which receives abundant rainfall and is strongly influenced by coastal and orographic processes, both CNN and LSTM models closely align with CHIRPS observations. However, CNN slightly outperforms LSTM in terms of correlation and normalized standard deviation. This advantage is attributed to CNN's superior spatial feature extraction, which captures fine-scale variability and gradients prevalent in high-precipitation zones.

In the Savannah zone, characterized by transitional rainfall patterns and active mesoscale convection, CNN again demonstrates stronger performance than LSTM and the raw GCMs. Both deep learning models show improved skill in representing rainfall variability, but CNN's ability to reconstruct spatial rainfall progression and zonal bands gives it a distinct edge.

Conversely, in the Sahel zone, where rainfall is sparse, erratic, and heavily influenced by interannual variability, LSTM exhibits marginally better performance than CNN. This is likely due to LSTM's recurrent architecture, which captures latent temporal dependencies and episodic precipitation events more effectively than the purely spatial CNN approach.

These findings suggest that CNN is better suited for wetter, spatially complex regions, while LSTM offers advantages in drier zones with pronounced temporal variability, reinforcing the need to tailor downscaling strategies to regional climatic dynamics.

In summary, the Taylor diagrams demonstrate that the CNN consistently outperforms ConvLSTM across all three climatic zones of West Africa. CNN achieves the best balance of accuracy metrics (standard deviation, correlation, and RMSE), followed closely by LSTM. The raw GCM exhibits the poorest skill in all zones. These findings underscore the superior spatial downscaling ability of CNN models in reproducing observed precipitation statistics under varied climatic regimes.

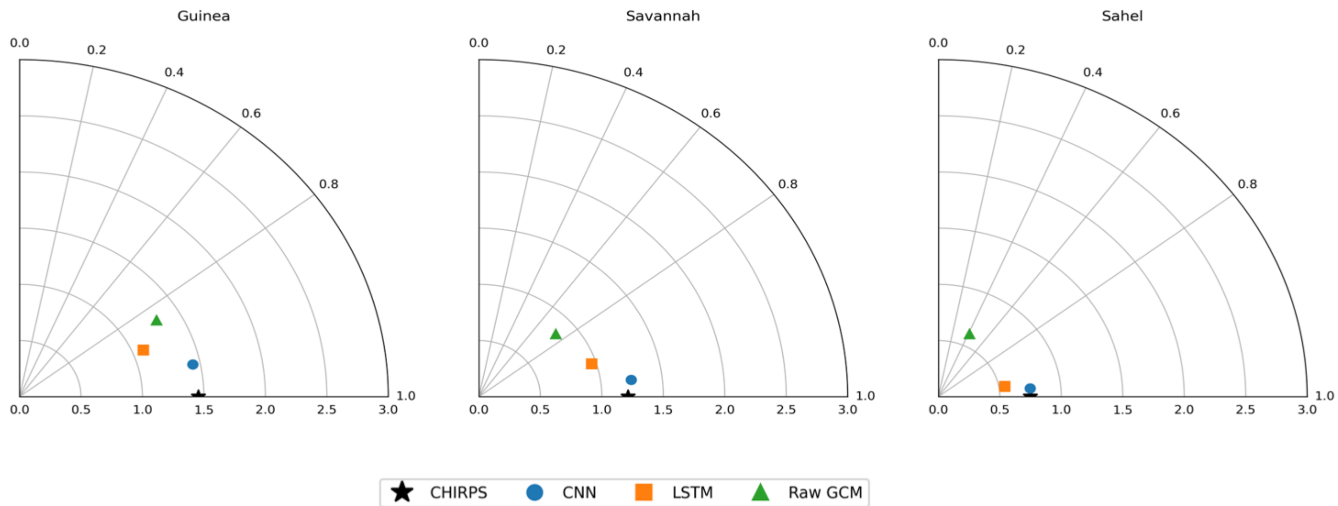



*Figure 59 Taylor Diagram compare the performance of the ConvLSTM and CNN for precipitation Over West Africa different climatic zone such as: Guinea (lat:4-8N), Savannah (lat:8-12N), and Sahel zone(lat:12-16N).*

### 3.4 Strengths of the Approach

The application of machine learning (ML) techniques specifically Convolutional Neural Networks (CNNs) and Convolutional Long Short-Term Memory (ConvLSTM) networks demonstrated several key advantages in enhancing the spatial and temporal fidelity of climate projections over West Africa:

First, improved spatial resolution was a central strength. Both CNN and LSTM models effectively refined the coarse outputs of Global Climate Models (GCMs), producing fine-scale climate fields that better represent regional gradients, local topographic influences, and eco-climatic transitions. This spatial enhancement is particularly valuable for applications such as localized vulnerability assessments, infrastructure planning, and agricultural zoning.

Second, the architectural adaptability of the models was evident across multiple CMIP6 GCMs—including CESM2-WACCM, FGOALS-g3, NorESM2-LM, and CMCC-CM2-SR5. The consistent performance of the downscaling models across structurally diverse climate models highlights their robustness and generalizability in regional climate modeling.

Third, the use of multi-model integration with three GCMs selected per two climate variable (temperature and precipitation) increased the robustness of the evaluation. This approach helped account for structural uncertainties inherent in climate models, as recommended by the Intergovernmental Panel on Climate Change (IPCC, 2021), and provided a more comprehensive basis for model intercomparison and performance benchmarking.

## 3.5 Limitations and Challenges

Despite the clear strengths, several limitations and methodological challenges remain:

First, model-specific biases persist. While CNN models excel at capturing spatial structure of precipitation, they often lack the ability to represent temporal dynamics, leading to limitations in simulating time-evolving phenomena such as monsoon progression or seasonal transitions. Conversely, LSTM-based models, while adept at temporal sequencing, may underrepresent spatial extremes particularly in the case of convective rainfall in arid or mountainous zones due to their smoothing tendencies.

Second, the approach is highly data-dependent. Training robust ML-based downscaling models requires access to high-resolution observational or reanalysis datasets such as ERA5 (for temperature) and CHIRPS (for precipitation). In data-scarce regions or when projecting into future periods where such observations are unavailable, model performance and reliability may be compromised.

Third, computational demand**s** pose a practical constraint. Deep learning architectures, especially ConvLSTM, require substantial computational resources for training, validation, and hyperparameter optimization. This may limit their operational scalability in low-resource settings or hinder iterative model development for ensemble-based applications.

Finally, while ML-based models demonstrate strong interpolation capacity, their ability to extrapolate under unprecedented future climate conditions remains limited. Without explicit integration of physical constraints or hybrid modeling frameworks, such models may struggle to maintain reliability when climate variables move outside the statistical bounds of historical training data (Vandal et al., 2019).

In summary, while CNN and LSTM-based statistical downscaling models offer transformative potential for regional climate modeling, careful attention must be paid to their inherent limitations, data dependencies, and computational requirements to ensure reliable and actionable outputs under future climate scenarios.

# CONCLUSION AND PERSPECTIVES

## 1. Conclusion

This study evaluated the application of machine learning (ML) and deep learning (DL) techniques specifically Convolutional Neural Networks (CNNs) and Convolutional Long Short-Term Memory (ConvLSTM) models for statistical downscaling of temperature and precipitation over West Africa. By leveraging these advanced models, the research aimed to enhance the spatial resolution and predictive accuracy of climate projections derived from coarse-resolution CMIP6 Global Climate Models (GCMs). The study focused on three GCMs per variable and covered both historical (1985–2014) and future (2015–2100) periods, thereby providing a robust assessment of downscaling performance across space and time.

In addressing Objective 1, the results demonstrate that CNN models significantly improved the spatial representation of temperature fields compared to raw GCM outputs. They effectively captured ecological and topographic gradients particularly in coastal, transitional, and orographic zones. However, ConvLSTM outperformed CNN in reproducing the temporal structure of temperature trends, achieving superior accuracy metrics such as lower RMSE and higher correlation with ERA5 observations. These findings full support Hypothesis 1, confirming CNN's strength in spatial refinement but highlighting LSTM’s superior temporal learning capacity for temperature modeling.

Regarding Objective 2, CNNs models exhibited a clear advantage in downscaling precipitation. They provided improved spatiotemporal coherence and more realistic representations of monsoonal dynamics, outperforming both ConvLSTMs and raw GCMs in capturing rainfall variability across West Africa. Notably, ConvLSTM downscaling enhanced rainfall realism in highland regions and river basins while reducing biases in key hydrological zones. These findings not support Hypothesis 2 as defined**,** but contrary show the capability of CNN architectures to model sequential climatic dependencies that are critical for precipitation forecasting.

From all analysis, both DL models substantially outperformed raw GCM outputs, reducing systematic biases and revealing localized climatic features critical for impact assessment and adaptation planning. The CNN model is more suited for applications requiring high spatial granularity, such as vulnerability mapping and land-use planning, while the ConvLSTM model is more appropriate for applications that depend on temporal consistency, such as hydrological modeling and early warning systems.

This research affirms the value of integrating data-driven downscaling frameworks into regional climate modeling efforts. By addressing structural limitations in GCM outputs and enhancing climate information at actionable scales, the study contributes to the growing body of knowledge supporting ML and DL approaches as viable tools for improving climate resilience in vulnerable regions such as West Africa.

## 2. Summary of Findings

This study demonstrates the substantial effectiveness of machine learning–based statistical downscaling approaches, particularly Convolutional Neural Networks (CNN) and Long Short-Term Memory networks (LSTM), in improving the accuracy and spatial resolution of climate projections over West Africa.

In the case of temperature, the LSTM model emerged as the most robust, achieving the lowest RMSE (0.212 °C) and MAE (0.156 °C), along with the highest $R^2$ (0.990) and near-perfect correlation ($r = 0.996$). These results underscore LSTM's ability to capture the temporal dynamics inherent in temperature variability, making it particularly well-suited for temperature downscaling. Although CNN also enhanced temperature projections relative to raw GCM outputs, it exhibited a slight cold bias (–0.729 °C) and marginally lower predictive accuracy. In contrast, the raw GCM consistently demonstrated the poorest performance ($R^2 = 0.653$), confirming the inadequacy of uncorrected GCM outputs for regional-scale applications.

For precipitation, the CNN model provided superior performance across most evaluation metrics, including a low RMSE (0.137 mm/day), high $R^2$ (0.995), and effective spatial reproduction of rainfall patterns, particularly in humid and orographic zones. The CNN model excelled in capturing both mean and extreme precipitation events like 95th percentile anomalies and outperformed LSTM in spatial consistency across West Africa. While the LSTM model maintained strong correlation ($r = 0.989$) and improved spatial representation over raw GCMs, it showed higher error margins and a stronger negative bias (–0.16 mm/day), especially in extreme rainfall zones. Raw GCM outputs remained the least reliable, characterized by widespread overestimation, coarse spatial resolution, and high RMSE (0.733 mm/day, $R^2 = 0.863$).

Taylor diagram evaluations further reinforced these findings, with CNN consistently outperforming other models in reproducing observed precipitation patterns across the Guinea, Savannah, and Sahel climatic zones. LSTM, while slightly less accurate in precipitation, delivered better performance in drier regions such as the Sahel due to its temporal generalization capacity. Overall, the comparative results suggest that CNN is better suited for spatial refinement of precipitation, whereas LSTM is more effective for temporally driven variables such as temperature. Both models significantly outperformed raw GCM outputs, offering clear evidence that deep learning based downscaling provides a reliable, high-resolution alternative for regional climate assessments and adaptation planning in West Africa.

## 3. Implications for Regional Climate Modeling

The findings of this study offer important implications for enhancing climate modeling and decision-making frameworks across West Africa: (1) Climate Adaptation Planning; the improved spatial fidelity of downscaled temperature and precipitation datasets enhances the reliability of localized climate information, thereby supporting more effective adaptation strategies in agriculture, water resource management, and urban infrastructure development. (2) Early Warning and Disaster Preparedness; the ability of ConvLSTM models to capture temporal dynamics provides a foundation for developing more accurate seasonal and sub-seasonal forecasts, which are critical for early warning systems related to droughts, floods, and heatwaves. (3) Model Selection for Sectoral Applications**;** the choice of downscaling model can be tailored to the nature of the climatic variable and its application. CNNs are particularly effective for spatially detailed precipitation modeling, while ConvLSTM models offer superior performance for temperature-related applications requiring temporal continuity such as public health planning, energy demand forecasting, and climate services. (4) Pathways for Future Development; the complementary strengths of CNN (spatial enhancement) and ConvLSTM (temporal modeling) highlight the potential of hybrid modeling approaches. Future research should explore integrated frameworks that combine these architectures to further improve the robustness and applicability of downscaled climate projections in data-scarce and climatically diverse regions like West Africa.

## 4. Perspectives for Future Work

While this study and results demonstrate the clear advantages of machine learning and deep learning in climate downscaling, several opportunities and future directions emerge: (1) Development of Hybrid Architectures*:*A promising avenue is the combination of CNN and ConvLSTM into hybrid models**,** which could simultaneously capture complex spatial and temporal patterns, especially for highly variable parameters like precipitation. (2) Incorporation of Physical Constraints*:* Integrating physical process knowledge into ML frameworks like physics-informed neural networks can improve generalization under changing climate conditions and reduce overfitting to observational datasets. (3) Extension to Future Climate Projections*:* While this study focused on historical performance, applying the trained models to downscale future GCM projections across different SSP scenarios will provide more actionable insights for climate change impact assessments and adaptation planning. (4) Operational Climate Services*:* With optimization, these ML-based models can be embedded into real-time or seasonal early warning system**s** for floods, droughts, and heat extremes enhancing resilience and preparedness across sectors. (5) Regional Transferability*:* Testing the models in other African subregions like Est, South, and Central Africa would help assess their transferability and guide broader deployment of ML-based downscaling tools in Africa and other data-scarce regions.

# BIBLIOGRAPHY REFERENCES


Abdullahi, J., Tahsin, A., Yesilnacar, M. I., Karabulut, A. İ., & Daramola, O. (2024). Assessment of Climate Change Impact on Precipitation Using Machine Learning Based Statistical Downscaling Method. In *E3S Web of Conferences* (Vol. 489, p. 04004). EDP Sciences.

Ahmed, K., et al. (2015). "Application of ANN in climate downscaling." Climate Dynamics, 44(3-4), 575–590.

Ayugi, B. O., Tan, G., Gao, C., Ngoma, H., & Ongoma, V. (2021). Evaluation of CMIP6 models in simulating precipitation and temperature over West Africa. *Atmosphere*, 12(4), 473. https://doi.org/10.3390/atmos12040473

Benestad, R. E., Parding, K. M., Dobler, A., & Mezghani, A. (2017). A strategy to effectively downscale climate projections. *Frontiers in Environmental Science*, 5, 1-9.

Bottou, L. (2010). "Large-scale machine learning with stochastic gradient descent." Journal of Machine Learning Research, 10, 1207–1233.

Breiman, L., et al. (1984). "Classification and Regression Trees." Wadsworth & Brooks/Cole.

Biasutti, M. (2019). "Rainfall trends in the African Sahel: Characteristics, processes, and causes." *WIREs Climate Change*, 10(4), e591.

Baño-Medina, J., Manzanas, R., & Gutiérrez, J. M. (2021). Configuration and intercomparison of deep learning neural models for statistical downscaling. *Environmental Modelling & Software*, 140, 105034. https://doi.org/10.1016/j.envsoft.2021.105034

Dosio, A., et al. (2015). Evaluation of RCM Precipitation Performance in West Africa. *Environmental Research Letters*, 10(12), 124006.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). "Overview of the CMIP6 experiment design." *Geoscientific Model Development*, 9(5), 1937–1958.

**Fouotsa Manfouo, N. C., Potgieter, L., Watson, A., & Nel, J. H. (2023).** A comparison of the statistical downscaling and long-short-term-memory artificial neural network models for long-term temperature and precipitation forecasting. *Atmosphere, 14*(4), 708. https://doi.org/10.3390/atmos14040708

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Rowland, J., ... & Michaelsen, J. (2015). The Climate Hazards Infrared Precipitation with Stations—A new environmental record for monitoring extremes. *Scientific Data*, 2, 150066. https://doi.org/10.1038/sdata.2015.66

Gaitan, C. F., Cannon, A. J., & Hsieh, W. W. (2014). Comparison of statistically downscaled temperature and precipitation fields derived from general circulation models. *International Journal of Climatology*, 34(10), 3061-3077. https://doi.org/10.1002/joc.3897

Giorgi, F., & Gutowski, W. J. (2015). Regional Dynamical Downscaling and the CORDEX Initiative. *Annual Review of Environment and Resources*, 40(1), 467–490. https://doi.org/10.1146/annurev-environ-102014-021217

**Goodarzi, M. R., Hashemipour, Z., Saremi, A., & Niazkar, M. (2024).** Intercomparison of machine learning methods for statistical downscaling of daily temperature under CMIP6 scenarios: A case study from Iran. *Journal of Hydroinformatics, 26*(5). https://doi.org/10.2166/hydro.2024.227

González-Abad, J., & Gutiérrez, J. M. (2024, November 6). *Are Deep Learning Methods Suitable for Downscaling Global Climate Projections? Review and intercomparison of existing models* [Preprint]. *arXiv*. https://doi.org/10.48550/arXiv.2411.05850

Gumindoga, W., Rientjes, T. H., Haile, A. T., Makurira, H., & Reggiani, P. (2019). Performance of bias-correction schemes for CMORPH rainfall estimates in the Zambezi River basin. *Hydrology and earth system sciences*, *23*(7), 2915-2938.

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). "Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset." *Scientific Data*, 7(1), 109.

Hewitson, B., et al. (2014). Regional climate information — Evaluation and projections. In: *Climate Change 2014: Impacts, Adaptation, and Vulnerability* (IPCC WGII AR5, Chapter 21).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... & Simmons, A. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. https://doi.org/10.1002/qj.3803

IPCC (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report. https://www.ipcc.ch/report/ar6/wg1

IPCC (2021). *Sixth Assessment Report (AR6), Working Group I: The Physical Science Basis*. Intergovernmental Panel on Climate Change.

Janicot, S., Omotosho, J. B., & Camberlin, P. (2021). West African Monsoon and climate change. In *The Climate of Africa* (pp. 189–218). Springer.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., & Joseph, D. (1996). "The NCEP/NCAR 40-year reanalysis project." *Bulletin of the American Meteorological Society*, 77(3), 437–471.

Ke, G., et al. (2017). "LightGBM: A highly efficient gradient boosting model." Advances in Neural Information Processing Systems, 30, 3146–3154.

Laprise, R. (2008). "Regional climate modeling: Benefits and limitations." Journal of Climate, 21(10), 2701–2722.

Lebel, T., Parker, D. J., Flamant, C., & Galle, S. (2003). "The AMMA field campaigns: Multiscale and multidisciplinary observations in the West African region." *Meteorology and Atmospheric Physics*, 95(3), 73–95.

Liu, K., Yao, C., Chen, J., Li, Z., Li, Q., & Sun, L. (2017). Comparison of three updating models for real time forecasting: A case study of flood forecasting at the middle reaches of the Huai River in East China. *Stochastic Environmental Research and Risk Assessment*, *31*, 1471-1484.

Maraun, D., Wetterhall, F., Ireson, A. M., et al. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, 48(3). https://doi.org/10.1029/2009RG000314

Maraun, D. (2016). "Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue." *Journal of Climate*, 29(7), 2545–2557.

Nicholson, S. E. (2013). "The West African Sahel: A review of recent studies on rainfall regimes and their interannual variability." *ISRN Meteorology*, 2013, 1–32.

Nikulin, G., et al. (2012). Precipitation Climatology in an Ensemble of CORDEX-Africa Regional Climate Simulations. *Journal of Climate*, 25(18), 6057–6078.

Omay, P.O., Muthama, N.J., Oludhe, C. *et al*. Projected changes in wet and dry extremes in the CMIP6 multi-model ensemble over the IGAD region of Eastern Africa. *Discov Atmos* **2**, 11 (2024). https://doi.org/10.1007/s44292-024-00016-4

Ogunjo, S., Olusola, A., & Olusegun, C. (2023). Predicting River Discharge in the Niger River Basin: A Deep Learning Approach. *Applied Sciences*, *14*(1), 12.

Oladeji, T. A., Adeyemo, B. T., & Awe, O. O. (2024). Predicting Precipitation Time Series Dynamics in West Africa Using Deep Learning Models. In *Practical Statistical Learning and Data Science Methods: Case Studies from LISA 2020 Global Network, USA* (pp. 111-125). Cham: Springer Nature Switzerland.

**Odekunle, T. O., & Eludoyin, A. O. (2008).** Sea surface temperature patterns in the Gulf of Guinea: their implications for the spatio-temporal variability of precipitation in West Africa. *International Journal of Climatology*, 28(11), 1507–1517

Pan, B., Hsu, K., AghaKouchak, A., & Sorooshian, S. (2019). Improving precipitation estimation using convolutional neural networks. *Water Resources Research*, 55(3), 2301–2321. https://doi.org/10.1029/2018WR024090

Polasky, A., Evans, J.L. & Fuentes, J.D. Statistical downscaling for precipitation projections in West Africa. *Theor Appl Climatol* **155**, 327–347 (2024). https://doi.org/10.1007/s00704-023-04637-4

Quenum, G.M.L.D., Klutse, N.A.B., Dieng, D. *et al.* Identification of Potential Drought Areas in West Africa Under Climate Change and Variability. *Earth Syst Environ* **3**, 429–444 (2019). https://doi.org/10.1007/s41748-019-00133-w

Rampal, N., Hobeichi, S., Gibson, P. B., Baño-Medina, J., Abramowitz, G., Beucler, T., González-Abad, J., Chapman, W., Harder, P., & Gutiérrez, J. M. (2024). Enhancing Regional Climate Downscaling through Advances in Machine Learning. *American Meteorological Society*. DOI: 10.1175/AIES-D-23-0066.1 https://doi.org/10.1175/AIES-D-23-0066.1

Semenov, M.A., & Barrow, E.M. (1997). "Use of a weather generator in climate impact studies." Agricultural and Forest Meteorology, 90(1-2), 1–12.

Šeparović, L., Alexandru, A., Laprise, R., Martynov, A., Sushama, L., Winger, K., ... & Valin, M. (2013). Present climate and climate change over North America as simulated by the fifth-generation Canadian regional climate model. *Climate Dynamics*, *41*(11), 3167-3201.

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., & Woo, W.-C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in Neural*

*Information Processing Systems (NeurIPS)*, 28. https://papers.nips.cc/paper_files/paper/2015/file/07563a3fe3bbe7e3ba84431ad9d055af-Paper.pdf

Sultan, B., & Janicot, S. (2003). "The West African monsoon dynamics. Part II: The 'pre-onset' and 'onset' of the summer monsoon." *Journal of Climate*, 16(21), 3407–3427.

Sylla, M. B., Nikiema, P. M., Gibba, P., Kebe, I., & Klutse, N. A. B. (2015). Climate Change over West Africa: Recent Trends and Future Projections. In *Adaptation to Climate Change and Variability in Rural West Africa* (pp. 25–40). Springer.

Sylla, M. B., et al. (2013). Climate Change Projections for West Africa with Regional Climate Models. *Climate Dynamics*, 40(7–8), 2137–2155.

Sylla, M. B., et al. (2010). Multimodel Ensemble Consistency of GCM-RCM Simulations over West Africa. *Theoretical and Applied Climatology*, 103, 385–400.

Tripathi, S., et al. (2006). "Downscaling precipitation using ANN and SVM models." Journal of Hydrology, 330(3-4), 621–640.

Vandal, T., et al. (2019). Quantifying uncertainty in discrete-continuous and spatially dependent data using deep learning. *Earth and Space Science*, 6(12), 2289–2301. https://doi.org/10.1029/2019EA000582

Vandal, T., Kodra, E., Ganguly, S., et al. (2019). DeepSD: Generating high-resolution climate change projections through single-image super-resolution. *Computers & Geosciences*, 126, 118–125. https://doi.org/10.1016/j.cageo.2019.02.004

Vandal, T., et al. (2019). Deep Learning for Climate Downscaling. *Scientific Reports*, 9(1), 15485.

Wilby, R. L., & Dawson, C. W. (2013). The statistical downscaling model: Insights from one decade of application. *International Journal of Climatology*, 33(7), 1707-1719. https://doi.org/10.1002/joc.3528

Zorita, E., & von Storch, H. (1999). The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate*, 12(8), 2474-2489.

Zhang, Y., et al. (2023). "Comparison of LSTM and SDSM for climate forecasting in Lake Chad Basin." Climate Research, 92(2), 145–159.

Zhang, Y., et al. (2021). "Comparative analysis of CMIP6 GCMs for West Africa climate modeling." *Climate Research*, 92(2), 145–159.