# Statistical forecasting of synoptic-scale rainfall in tropical Africa

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
(Dr. rer. nat)
von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

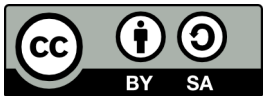
von

**M.Sc. Athul Rasheeda Satheesh**

aus Thiruvananthapuram

| | |
|---|---|
| Tag der mündlichen Prüfung: | 12.01.2024 |
| Referent: | Prof. Dr. Peter Knippertz |
| Korreferent: | Prof. Dr. Andreas H. Fink |

Karlsruher Institut für Technologie



# Abstract

Numerical model-based rainfall forecasts consistently fail to outperform more simplistic climatology-based forecasts over tropical Africa. The limitations of contemporary numerical weather prediction (NWP) models in resolving small-scale convective processes and the decreased intrinsic predictability in tropical areas are the major reasons for this. However, current research suggests that data-driven forecasts, which rely exclusively on spatio-temporal dependencies derived from satellite-based gridded rainfall estimates, demonstrate potential for improved predictability of synoptic-scale rainfall over tropical Africa. A further benefit is the lower cost associated with statistical models for producing such forecasts. Motivated by this, the primary goal of this dissertation is to develop daily rainfall forecasts using statistical models trained solely on predictor variables derived from observational datasets.

First, we investigate the potential for constructing a statistical model to forecast the probability of daily precipitation occurrence across tropical Africa utilizing a logistic regression (Logistic) model. We apply the newly developed Coefficient of Predictive Ability (CPA) to detect predictors based on rainfall data from the previous three days. The Logistic model is trained on the identified predictors in a 'leave-one-out' cross-validation mode to improve the robustness of output. The statistical forecast is compared with a climatology-based benchmark, the operational ensemble precipitation forecast (ENS) from the European Center for Medium Weather Forecast (ECMWF) and a statistically postprocessed ensemble forecast. All forecasts show poor skill within the main rainbelt region, with statistically insignificant differences in the Brier skill score (BSS). However, the Logistic forecast outperforms the benchmarks mainly along the fringes of the rainbelt. The CPA, along with a coherent-linear-propagation factor (coh) developed in the thesis and auxiliary meteorological variables, indicate that the high stochasticity in the rainbelt limits predictability. Along the fringes of the rainbelt, however, the Logistic model leverages coherently propagating precipitation features associated with potent tropical weather systems like African easterly waves (AEWs) in the summer season.

Secondly, this thesis examines the impact of tropical waves on synoptic-scale rainfall and explores the influence of their phases on ENS forecast. The analysis indicates that fast-moving waves, like the tropical depressions (TDs) and Kelvin waves, have the most significant impact on daily rainfall anomalies over tropical Africa. Slow, large-scale

waves, such as the Madden-Julian Oscillation (MJO) and equatorial Rossby (ER) waves, demonstrate some modulation, but the corresponding propagation behaviour does not show expected wave structures in space. The phases of all tropical waves apart from the eastward inertio-gravity (EIG) waves have a discernible impact on the skill of the ENS forecast. However, the influence is more noticeable in predicting the occurrence of rainfall than in estimating the amount of precipitation.

One major drawback of the data-driven approach described earlier is its reliance on human expertise when choosing predictors. Existing knowledge of coherent wave features is not included directly, specifically in regions where multiple wave types may be active, which can result in non-linear interactions potentially undetected by CPA. However, a more significant limitation is the use of a logistic regression model, which can only predict the likelihood of rainfall occurrence. Further statistical models employed in this thesis address these two limitations. To overcome the former, a gradient-boosting technique has been implemented to partially automate predictor selection. As the logistic regression model is constrained to binary outcomes, a gamma regression model and a more advanced convolutional neural network (CNN) model have been utilised to generate six-hour-ahead deterministic forecasts of daily precipitation amounts. The recently introduced Easy Uncertainty Quantification (EasyUQ) tool transforms these deterministic forecasts into readily calibrated probabilistic forecasts. Predictors based on the convectively coupled tropical waves used earlier are identified to train the statistical models in a 'leave-one-out' cross-validation mode. The statistical forecasts produced are compared with various benchmark forecasts, including two climatology-based forecasts, the ENS forecast and a probabilistic forecast generated through EasyUQ from the ECMWF control forecast (ENS-control EasyUQ). Overall, the statistical forecasts demonstrate proficiency throughout the entire analysis domain, attesting to the robustness of the tropical wave-based predictors. In the rainbelt region, the statistical forecasts are more skilful than the benchmark forecasts. The differences between the gamma regression and the benchmark forecasts demonstrate statistical significance over a wider area than the CNN forecast. In regions with very low rainfall amounts, specifically the northern Sahel and the equatorial Atlantic Ocean, the ENS-control EasyUQ prediction appears to display an enhanced level of skill. However, differences between the statistical and benchmark forecasts lack statistical significance in most such regions.

This dissertation belongs to the first studies to systematically present the advantages of employing statistical forecasts to predict daily rainfall in tropical Africa. Based on meteorological evidence, this thesis attempts to provide insights into the underlying principles utilised by complex machine-learning models in generating more skilful forecasts than the ones produced by conventional NWP models. The statistical models discussed in this dissertation are developed with a focus on operational forecasting and can be deployed with minimal changes due to lower costs compared to their NWP counterparts.

# Kurzfassung

Niederschlagsvorhersagen über dem tropischen Afrika basierend auf numerischen Modellen erreichen selten eine höhere Güte als einfachere klimatologische Vorhersagen. Die Grenzen der heutigen numerischen Wettervorhersagemodelle (NWP) bei der Auflösung kleinräumiger Konvektionsprozesse und die geringere intrinsische Vorhersagbarkeit in tropischen Gebieten sind wesentliche Gründe dafür. Aktuelle Forschungsergebnisse deuten jedoch darauf hin, dass datengestützte Vorhersagen, die sich ausschließlich auf räumlich-zeitliche Abhängigkeiten stützen, ein Potenzial für eine verbesserte Vorhersagbarkeit von Niederschlägen auf synoptischer Ebene über dem tropischen Afrika aufweisen. Ein weiterer Vorteil sind die geringeren Rechenkosten, die mit der Erstellung statistischer Modelle für solche Vorhersagen verbunden sind. Das Hauptziel dieser Dissertation besteht darin, tägliche Niederschlagsvorhersagen mit statistischen Modellen zu entwickeln, die ausschließlich mittels Vorhersagevariablen aus Beobachtungsdatensätzen trainiert werden.

Zunächst untersuchen wir das Potenzial für die Konstruktion eines statistischen Modells zur Vorhersage der Auftrittswahrscheinlichkeit von täglichem Niederschlag im tropischen Afrika unter Verwendung eines logistischen Regressionsmodells (Logistic). Wir wenden den neu entwickelten ‘Coefficient of Predictive Ability’ (CPA) an, um Prädiktoren auf der Grundlage der Niederschlagsdaten der vorangegangenen drei Tage zu ermitteln. Das logistische Modell wird anhand der identifizierten Prädiktoren in einem "Leave-one-out"-Kreuzvalidierungsmodus trainiert, um die Robustheit der Ergebnisse zu verbessern. Die statistische Vorhersage wird mit einem klimatologie–basierten Benchmark, der operationellen Ensemble-Niederschlagsvorhersagen (ENS) des Europäischen Zentrums für mittelfristige Wettervorhersage (ECMWF) und einer statistisch nachbearbeiteten Ensemble-Vorhersage verglichen. Alle Vorhersagen weisen innerhalb des tropischen Regengürtels eine geringe Güte auf, mit statistisch nicht-signifikanten Unterschieden im Brier Skill Score (BSS). Die logistische Vorhersage übertrifft jedoch die Benchmarks hauptsächlich an den Rändern des Regengürtels. Die CPA, zusammen mit dem neuartigen ‘coherent-linear-propagation factor’ (coh) und meteorologischen Hilfsvariablen, zeigen, dass die hohe Stochastizität im Regengürtel die Vorhersagbarkeit einschränkt. Entlang der Ränder des Regengürtels nutzt das logistische Modell jedoch sich kohärent ausbreitende Niederschlagsmerkmale, die mit starken tropischen Wettersystemen wie den ‘African easterly waves’ (AEWs) im Sommer der Nordhemisphäre verbunden sind.

Zweitens, werden in dieser Studie die Auswirkungen der verschiedenen Phasen tropischer Wellen auf synoptisch-skalige Niederschläge untersucht. Die Studie untersucht auch den Einfluss der Phasen auf NWP- und statistische Niederschlagsvorhersagen. Die Analyse zeigt, dass schnell propagierende Wellen, wie die tropische Depressionen (TDs) und Kelvin-Wellen, den größten Einfluss auf die täglichen Niederschlagsanomalien über dem tropischen Afrika haben. Großräumige Wellen, wie die Madden-Julian-Oszillation (MJO) und die äquatorialen Rossby-Wellen (ER), zeigen zwar eine gewisse Modulation, aber mit fehlenden Wellenstrukturen im verbundenen Niederschlagsmuster. Die Phasen aller tropischen Wellen, mit Ausnahme der ostwärts propagierenden Trägheitsschwerewellen (EIG) haben einen erkennbaren Einfluss auf die Güte der ENS-Vorhersage. Allerdings ist der Einfluss bei der Vorhersage des Auftretens von Niederschlägen stärker ausgeprägt als bei der Schätzung der Niederschlagsmenge.

Ein großer Nachteil des oben beschriebenen datengetriebenen Ansatzes besteht darin, dass er bei der Auswahl der Prädiktoren vollständig auf menschliches Fachwissen angewiesen ist. Vorhandenes Wissen über kohärente Wellenmerkmale wird nicht direkt einbezogen, insbesondere in Regionen, in denen mehrere Wellentypen aktiv sein können. Dies führt zu nichtlinearen Wechselwirkungen, die von der CPA möglicherweise nicht erkannt werden. Eine noch wichtigere Einschränkung ist jedoch das logistische Regressionsmodell, das lediglich die Wahrscheinlichkeit des Auftretens von Niederschlägen vorhersagen kann. Die Weiterentwicklung der statistischen Modelle soll diese beiden Mängel beheben. Um den ersten zu beheben, wurde ein Gradient-Boosting-Verfahren eingeführt, um die Auswahl der Prädiktoren teilweise zu automatisieren. Da das logistische Regressionsmodell auf binäre Ergebnisse beschränkt ist, wurden ein Gamma-Regressionsmodell und ein fortschrittlicheres Convolutional Neural Network (CNN) verwendet, um deterministische Vorhersagen der täglichen Niederschlagsmengen für sechs Stunden im Voraus zu erstellen. Das kürzlich entwickelte 'Easy Uncertainty Quantification' (EasyUQ)-Tool wird verwendet, um diese deterministischen Vorhersagen in leicht zu kalibrierende probabilistische Vorhersagen umzuwandeln. Es werden Prädiktoren auf der Grundlage von sieben konvektiv gekoppelten tropischen Wellen identifiziert, um die statistischen Modelle in einem "Leave-one-out"-Kreuzvalidierungsmodus zu trainieren. Die erstellten statistischen Vorhersagen werden mit verschiedenen Benchmark-Vorhersagen verglichen, darunter zwei klimatologie-basierte Vorhersagen, die rohe ECMWF-Ensemblevorhersage und eine probabilistische Vorhersage, die mit EasyUQ aus der ECMWF-Kontrollvorhersage (ENS-control EasyUQ) erstellt wurde. Insgesamt weisen die statistischen Vorhersagen im gesamten Analysebereich eine hohe Güte auf, was die Robustheit der auf tropischen Wellen basierenden Vorhersagen belegt. Im Bereich des tropischen Regengürtels sind die statistischen Vorhersagen besser als die Benchmark-Vorhersagen. Allerdings zeigt die Gamma-Regressionsvorhersage statistische Signifikanz über ein größeres Gebiet als die CNN-Vorhersage. In Regionen mit geringem Niederschlag, insbesondere in der nördlichen Sahelzone und im äquatorialen Atlantik, scheint die ENS-Kontrollvorhersage EasyUQ ein höhere Güte aufzuweisen. Allerdings

sind die Unterschiede zu den statistischen Vorhersagen in den meisten dieser Regionen statistisch nicht signifikant.

Diese Dissertation gehört zu den ersten Studien, die systematisch die Vorteile des Einsatzes statistischer Prognosen zur Vorhersage des täglichen Niederschlags im tropischen Afrika aufzeigen. Basierend auf meteorologischen Erkenntnissen versucht diese Arbeit, wertvolle Einblicke in die zugrundeliegenden Prinzipien, die von komplexen maschinellen Lernmodellen genutzt werden, zu liefern, um bessere Vorhersagen als die von konventionellen NWP-Modellen zu erstellen. Die in dieser Dissertation diskutierten statistischen Modelle wurden mit Schwerpunkt auf der operationellen Vorhersage entwickelt und können aufgrund ihrer im Vergleich zu ihren Pendants im NWP-Bereich wesentlich geringeren Kosten mit minimalen Änderungen eingesetzt werden.

# Preface

The PhD candidate confirms that the research presented in this thesis contains significant scientific contributions by himself. This thesis reuses material from the following publications:

> Rasheeda Satheesh, A., Knippertz, P., Fink, A.H., Walz, E.-M. & Gneiting, T. (2023): Sources of predictability of synoptic-scale rainfall during the West African summer monsoon. Quarterly Journal of the Royal Meteorological Society, 1–17. Available from: https://doi.org/10.1002/qj.4581

> Rasheeda Satheesh, A., Knippertz, P., Fink, A.H.(2024): Machine learning models for daily rainfall forecasting in Northern Tropical Africa using tropical wave predictors (in prep.).

The abstract and Chapters 1, 2, 3, 4, 5, 6, 9 and 10 reuse material from Rasheeda Satheesh et al. (2023) ©2023 The Authors, Quarterly Journal of the Royal Meteorological Society published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

The abstract and Chapters 1, 2, 3, 4, 5, 7, 8, 9 and 10 reuse material from Rasheeda Satheesh et al. (2024, in prep.).

The research leading to these results has been done within project C2 of the Transregional Collaborative Research Center SFB/TRR 165 “Waves to Weather”, funded by the German Research Foundation (DFG). The research proposal for this project was written by Peter Knippertz, Andreas H. Fink and Tilmann Gneiting. Analyses in Rasheeda Satheesh et al. (2023) were solely performed by the candidate, who also wrote the text with advice from Peter Knippertz, Andreas H. Fink and Tilmann Gneiting. The analyses in Chapters 7 and 8 were solely performed by the candidate with advice from Peter Knippertz and Andreas H. Fink.

The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others. This copy has been supplied on the understanding that this is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

# Contents

# 1. Introduction

The potential for systematic weather forecasts was considered as early as the turn of the twentieth century (Bauer et al., 2015). Research articles such as Abbe (1901) and Bjerknes (1904) popularised the idea that numerical weather prediction is fundamentally no different from any other initial value problem. The concept that the future state of the atmosphere can be predicted through solving the governing partial differential equations, although theoretically sound, was immensely ambitious during the given period as routine observations were scarcely available, and the computational resources required to carry out such an endeavour were still several decades away.

The creation of such a numerical model to predict the weather (now referred to as a numerical weather prediction model, or NWP model) was first attempted in Princeton, New Jersey, in 1932. It was not until 1958 that a fully functioning model was successfully deployed without any major defects (Shuman, 1989). With the deployment of satellites and the subsequent availability of satellite-based observations on a global scale by the end of the 1970s, the problem of sparse observations gradually reduced. Nonetheless, the predominant challenge that emerged in parallel was the management of the ever-growing volumes of data. Today, decades later, with the help of some of the world's most advanced computers and storage devices, many operational weather services regularly issue weather forecasts at different spatio-temporal scales. Investing in weather forecasting facilities is also widely accepted due to their ability to provide accurate forecasts, ultimately enabling the prevention of loss of life and property (Lazo et al., 2009).

However, various studies (e.g., Haiden et al., 2012; Bauer et al., 2015) present compelling evidence that, despite the above-mentioned advancements, precipitation forecasts in the tropics have not yet reached the level of accuracy desired, particularly in regions such as tropical Africa. Furthermore, the scarcity of in-situ observation is compounded by the reduced inherent predictability of synoptic-scale weather in tropical Africa as a result of faster error growth in comparison to midlatitude regions (Judt, 2020). Tropical Africa stands out as the sole region globally displaying negative values in the autocorrelation of daily precipitation amounts at a one-day lag (Roehrig et al., 2013). Studies, such as Vogel et al. (2018) and **?**, demonstrate that the latest synoptic-scale ensemble precipitation forecasts encounter difficulties in outperforming much simpler climatology-based forecasts, even after statistical correction for biases and dispersion

errors. The deficient forecasts may have severe implications not only for agriculture, a crucial source of income in tropical Africa, but also for flood and drought warning systems, reservoir management, and infrastructure maintenance, to mention a few. This potentially presents a significant threat to the well-being and the livelihoods of more than one billion people (United Nations, Department of Economic and Social Affairs, Population Division, 2022).

The presence of strong synoptic-scale forcing can lead to a coherent modulation of rainfall patterns, which may allow for potential windows of improved predictability. For example, Kniffka et al. (2020) reports an improvement in model skill in the Gulf of Guinea and neighbouring coastal areas during the summer months when synoptic-scale vortices are present. Numerous studies (e.g., Wheeler and Kiladis, 1999; Yang et al., 2003; Kiladis et al., 2009) highlight the significance of tropical waves (TWs) in synoptic-scale variability within tropical regions, particularly over tropical Africa (e.g., Reed et al., 1977; Fink and Reiner, 2003). Furthermore, African easterly waves (AEWs) and other tropical waves have the potential to influence synoptic-scale rainfall in northern tropical Africa, as concluded by various studies (Fink and Reiner, 2003; Lavaysse et al., 2006; Schlueter et al., 2019a,b). TWs occur at various temporal and spatial scales and have wavelengths ranging from several hundred to several thousand kilometres and time periods ranging from days to weeks. TWs in the lower troposphere also couple with convection and are known as Convectively Coupled Tropical (or Equatorial) Waves (CCTWs or CCEWs) (Wheeler and Kiladis, 1999; Kiladis et al., 2009). Dias et al. (2018) used precipitation forecasts based on the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) and the European Centre for Medium Weather Forecast (ECMWF) Integrated Forecast System (IFS) to examine the relationship between improved representation of TWs and tropical rainfall at synoptic scales. This study disclosed that both models were able to initialise and propagate disturbances in the form of tropical waves and that overall, the precipitation forecast skill in the tropics, although much lower compared to mid-latitudes, remained reasonable for about four days. However, no causal relationship between improved TW representation and improved rainfall forecasts was confirmed as this study was limited to just over a year of data due to the enormous computational cost of running NWP models.

This issue was subsequently addressed from a data-driven perspective (i.e., using a statistical model) by Li and Stechmann (2020), who used daily observational data and assumed a damped oscillator form for the TWs. Statistical models like this provide an alternative approach to weather forecasting, distinct from NWP models. They have been in existence long before the widespread adoption of NWP models (Gates and Tong, 1976; Gabriel and Neumann, 1962), but were not as popular as NWP models. This was partly because the success of such models does not explicitly rely on the physics of the atmosphere but rather on the coherent relationships between rainfall and other meteorological variables. However, they have been gradually gaining acceptance in the meteorological community due to advancements in statistical computation and, more

recently, the progress in the fields of machine learning (ML) and artificial intelligence (AI) as shown by Pathak et al. (2022), Lam et al. (2022) and Bi et al. (2022). Although these 'general' statistical models perform adequately compared to NWP models in terms of various meteorological variables and are comparatively much cheaper to operate, they face significant challenges in predicting tropical rainfall accurately.

Therefore, the primary aim of this dissertation is to construct reliable statistical models that can accurately predict daily rainfall in tropical Africa. These models will use suitable predictors which modulate rainfall at the synoptic-mesoscale, identified exclusively from observational datasets. Moreover, the outcomes of the statistical models will be supported with meteorological evidence where possible. To achieve this aim, the dissertation is organised as follows: Chapter 2 will establish the theoretical background pertaining to rainfall in tropical Africa along with the theory of tropical waves. It will also cover the current status of rainfall prediction in tropical Africa. Following this, the motivations and the specific research questions that need to be answered are introduced in Chapter 3. The details of the data utilized for this investigation and the methods applied in the analysis are expounded upon extensively in Chapter 4. Chapter 5 describes the statistical models and the novel diagnostics introduced or developed in this thesis. Chapter 6 aims to develop a simple statistical model for predicting daily rainfall occurrence in tropical Africa. Its capabilities are subsequently evaluated against various benchmarks, and its performance will be explained on the basis of meteorological evidence. Chapter 7 first identifies the regions with the most intense activity for each tropical wave before assessing potential relationships between the performance of rainfall forecasts based on NWP models and the phases of tropical waves calculated in the identified regions. The forecast skill is compared in terms of both rainfall occurrence and rainfall amounts. Chapter 8 discusses the development of more sophisticated machine learning (ML)–based models that aim to predict daily rainfall accumulation using predictor variables based exclusively on convectively coupled tropical waves used in Chapter 7. These predictions are also benchmarked, and, where possible, meteorological explanations are provided. Chapter 9 presents the major findings of the studies described in this thesis by addressing the key questions posed in Chapter 3. The final chapter provides a brief outlook and discusses the relevance of the results of this thesis in the context of advancing research to improve statistical rainfall prediction in tropical Africa.

# 2. Theoretical background

This chapter provides a detailed account of the theoretical background necessary for a thorough comprehension of this dissertation. The initial section focuses on the mean state of rainfall in Tropical Africa, briefly describes the West African Monsoon system, its seasonal variations influencing rainfall variability, and the major sources of rainfall variability over tropical Africa. The subsequent section outlines the theory and thoroughly examines the identification of tropical waves. The final section offers an outline of the evolution and present state of rainfall prediction in tropical Africa.

## 2.1 Rainfall in tropical Africa

### 2.1.1 Mean state

Tropical Africa represents one of Africa's most populous regions, with the local population heavily relying on the annual climate for their livelihood (Fink et al., 2017). Rainfed agriculture stands as one of the major sources of income in these regions. Precipitation patterns in tropical Africa exhibit substantial fluctuations and are season-dependent. While certain areas, such as the Sahel, rely on the lone rainy spell linked to the West African Monsoon (WAM, see Section 2.1.2), other regions, like the southern coastal regions of West Africa (the Guinea coast), depend on the double wet spells generated by the northward and southward shifts of the climatological rainbelt.

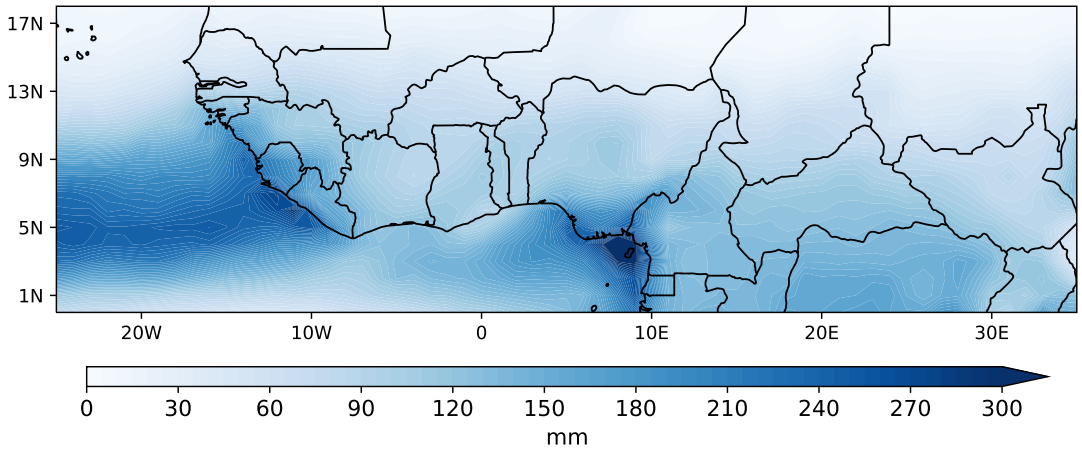



**Figure 2.1:** Monthly accumulated rainfall over tropical Africa from 2001-2019.

Orography is another significant factor that affects rainfall in tropical Africa. (Fig. 5.2). The terrain across tropical Africa exhibits significant diversity from area to area, ranging from highlands in Guinea, Cameroon and Ethiopia to lowlands covering the extensive plains of the Sahel. Fig. 2.1 illustrates the mean monthly accumulated rainfall over a period of 19 years, from 2001 to 2019. Some of the inferences that can be drawn are: (1) The drier north and the wetter south, (2) The anomalously dry area, referred to as the Ghana-Benin dry zone (Vollmert et al., 2003), stretching from Ghana to approximately central Benin, (3) The extremely moist regions situated along the western coastline in proximity to the Guinea mountains and the Cameroon highlands. (4) The moderately wet region encompassing the Congo Basin, Gabon, and Equatorial Guinea.

### 2.1.2 West African monsoon

Historically, monsoon systems are defined as the seasonal reversal of low-level winds that typically bring moist air from the ocean into the continent as a response to the heat-low generated over the continent. However, various regions across tropical Africa are classified as having monsoon climates, despite not all having a single wet season (for example, the Guinea coast) (Cook and Vizy, 2019). Therefore, in this dissertation, wet seasons in all the regions over tropical Africa will be collectively considered as the West African Monsoon.

The WAM plays a vital role in providing the bulk of rainfall in tropical Africa, with irregularities posing a potential threat to the well-being of over 300 million individuals (Raj et al., 2019). The onset of the WAM varies annually (Bello, 1997). It typically commences around late April to early May as the Saharan heat low (SHL) strengthens, drawing moist cold air into the continent from the equatorial Atlantic Ocean along the Guinea coast before intensifying during May and reaching its peak around late June to mid-July, as shown in Fig. 2.2 (Lavaysse et al., 2006; Fink et al., 2017). Between early June and early July, there is a sudden northward shift in the WAM, which results in the weakening of rainfall over the Guinea coast and a gradual strengthening over the Sahel region (see orange line in Fig. 2.2). The West African monsoon jump, as it is known, has been documented in numerous studies (e.g., Sultan and Janicot, 2000; Peyrille et al., 2016; Sultan and Janicot, 2003) and is thought to be caused by the strengthening and northward displacement of the African easterly jet (AEJ), which in turn gives rise to the AEWs (see Section 2.1.4.1).

The transition from the coast of Guinea to the Sahel region characterises the dynamic onset of the WAM, which contributes substantially to rainfall in the Sahel. The WAM then gradually strengthens and reaches its peak in mid-August, while the Guinea coast experiences the so-called 'Little Dry Season'. In late August to mid-September, the West African Monsoon weakens in the Sahel, causing the rainbelt to move towards the equator, setting in motion the second rainy season along the Guinea coast (refer to the blue line in Figure 2.2). As shown in Figure 2.2, it can be observed that the second rainy season along the Guinea coast is typically less intense compared to the first one.

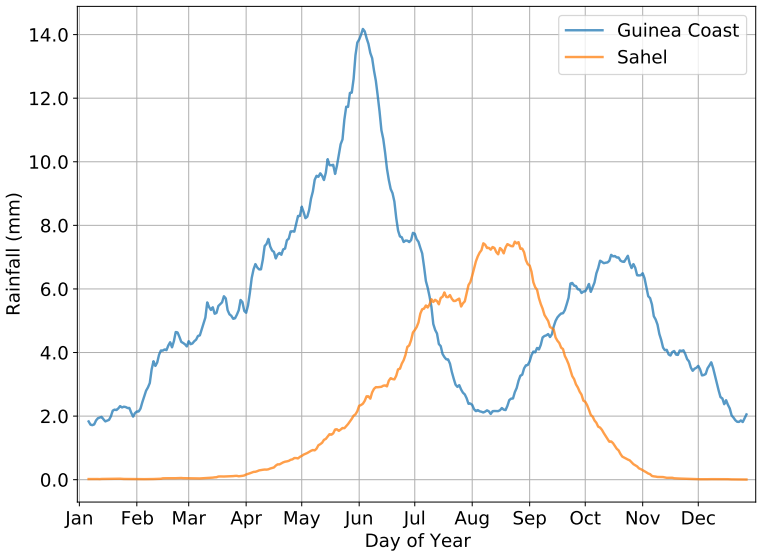



**Figure 2.2:** Area-averaged daily rainfall climatology from 2001–2019 over the Guinea Coast (4° N–6° N) and the Sahel (12° N–15° N). Both are computed over 12° W–6° E and smoothed using an 11-day running mean filter. Modified figure from Cook and Vizy (2019)

.

Unlike the first, the second rainy season shows a rather gradual decline towards the end of October after reaching its peak in the early to mid-October period. This marks the end of the West African monsoon.

### 2.1.3 Seasonal cycle

Rainfall in tropical Africa displays significant seasonal variability, as noted in Section 2.1.1. This variability is illustrated in Figure 2.3. During the winter months or DJF season (Fig. 2.3a), the highest amount of rainfall (climatological rainbelt) can be found over the Gulf of Guinea, while the majority of land areas, with the exception of the Congo basin, experience dry conditions. However, according to Knippertz and Fink (2009), rare precipitation events resulting from tropical-extratropical interactions, specifically tropical plumes (see section 2.1.4.2), in the northern regions of tropical Africa may occur during this time of the year. During the spring season (March-April, Fig. 2.3b), the climatological rainbelt moves northward, marking the commencement of the first rainy season in the southern coastal regions of West Africa.

The rainfall over the Guinea coast intensifies during the pre-monsoon season (MJ, Fig. 2.3c) as the rainbelt moves further north. While the African easterly jet (Fink and Reiner, 2003; Burpee, 1974; Lafore et al., 2017) strengthens over the Sahel region, weak African easterly waves propagate westward. At the same time, slow westward propagating cyclonic vortices (Knippertz et al., 2017) form over equatorial Africa, leading to abundant rainfall over the Gulf of Guinea. Rainfall in the far northern Sahel is still influenced by extratropical forcings (see the section, 2.1.4.2), specifically cold air surges

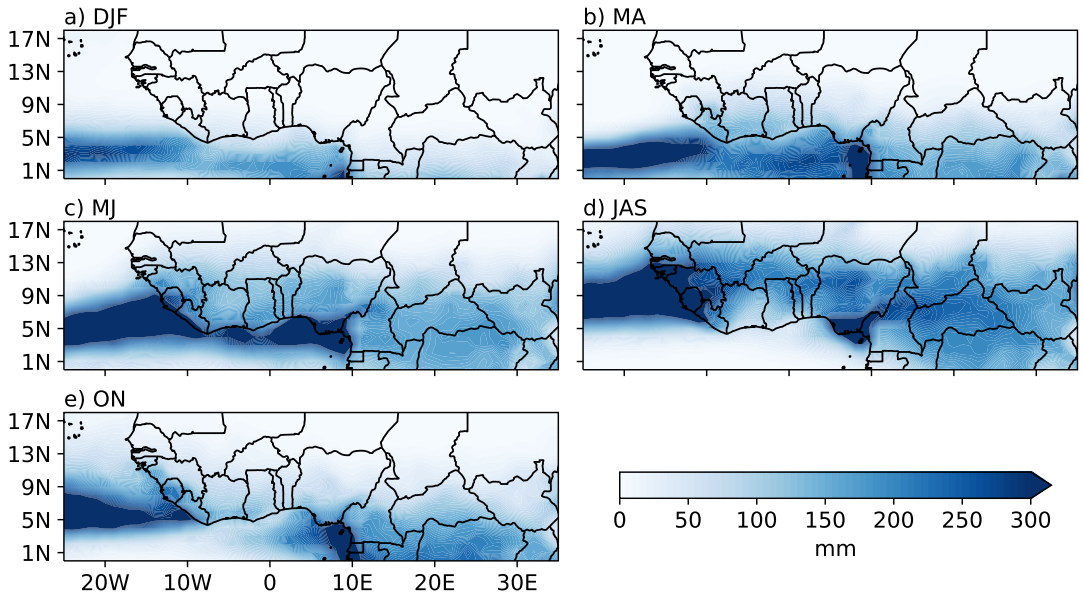



**Figure 2.3:** Monthly accumulated rainfall during a) December–February (DJF), b) March–April (MA), c) May–June (MJ), d) July–September (JAS), and e) October-November (ON) over tropical Africa from 2001-2019.

from the Mediterranean (Vizy and Cook, 2009, 2014). These forcings enhance convective activity over the western Sahel while suppressing it over the eastern Sahel.

Upon the dynamic onset of the WAM in JAS, the Sahel region receives copious amounts of rainfall in the form of well-organized, large-scale convective systems (Fink et al., 2017). The rainbelt, now situated slightly to the south of the Sahel and the central African region, is typically highly stochastic, and precipitation often (but not always) occurs as a result of the so-called 'popcorn convection'. In JAS (specifically during July), the rainbelt is situated at its northernmost position. This results in a brief hiatus of precipitation events along the Guinea coast, known as the Little Dry Season (Hamilton et al., 1945). However, slow-cyclonic vortex-driven precipitation persists over the Gulf of Guinea (Rasheeda Satheesh et al., 2023).

During ON, the WAM weakens, and the rainbelt shifts equatorward, initiating the second wet season along the Guinea coast. While weak AEWs (Fink and Reiner, 2003; Burpee, 1974; Lafore et al., 2017) may trigger and sustain precipitation events over the Sahel, cyclonic vortices (Knippertz et al., 2017) may bring about rainfall over the Gulf of Guinea.

### 2.1.4 Sources of variability

While various factors contribute to the intensity, location and duration of rainfall in tropical Africa during different seasons, those that are relevant to this thesis are described below.

#### 2.1.4.1 African easterly waves

AEWs were first identified in the 1930s by Piersig (1936) and Regula (1936). Later studies in the 1960s, such as Carlson (1969), investigated the formation of easterly wave disturbances over Africa that transformed into Atlantic hurricanes. Around the same time, Frank (1969) also investigated the connection between easterly waves and characteristic 'inverted V-shaped' clouds across Africa. Today, AEWs are known to be the strongest synoptic-scale feature in tropical Africa, contributing to about a third of the rainfall variability (Dickinson and Molinari, 2000; Lavaysse et al., 2006; Mekonnen et al., 2006). As their name suggests, they are a westward-propagating wave mode with phase speeds of around 8–10 $ms^{-1}$ (Reed et al., 1977) and a period of approximately 3–5 days (Burpee, 1972). Their zonal wavelength is about 4000 km (Kiladis et al., 2006; Diedhiou et al., 1999), although previous studies have estimated smaller wavelengths of around 2800 km (e.g., Burpee, 1975; Reed et al., 1977), indicating some degree of variability. AEWs are thought to form due to barotropic–baroclinic instabilities related to the African easterly jet (Burpee, 1972; Thorncroft and Hoskins, 1994). Although weaker AEWs may be detected over tropical Africa as early as May and as late as October (Burpee, 1974; Lafore et al., 2017), they are mainly active during July–September, when the AEJ is at its strongest due to the large temperature gradient between the hot Sahara and cool tropical Africa. However, a coherent model incorporating all these 'flavours' of AEWs is still lacking (Schlüter, 2019).

AEWs are also known to interact with other tropical waves. Previous research conducted by Ventrice and Thorncroft (2013) and, more recently, Lawton et al. (2022) have examined the influence of convectively coupled Kelvin waves in triggering and modulating AEWs. Another study by Jung and Knippertz (2023) found that convectively coupled westward inertio-gravity waves (WIGs, see section, 2.2) identified through fast-Fourier transform-based methods (see the section, 4.2 for details) may actually be fast-moving AEWs, indicating possible overlap between AEWs and WIGs in Fourier-space. Furthermore, Fink and Reiner (2003) estimate that roughly two-thirds of squall lines observed in West Africa can be attributed to the presence of AEWs. Since WIGs are frequently misidentified as squall lines (Tulich and Kiladis, 2012), some of these may serve as evidence of coupling between AEWs and WIGs. Finally, a recent study by Cheng et al. (2019) details an AEW-MRG hybrid mode, so named due to its combination of properties from both African Easterly Waves and mixed Rossby-gravity waves, indicating the interaction between the two waves.

#### 2.1.4.2 Tropical–extratropical interactions

Tropical-extratropical interactions, such as tropical plumes (e.g., Mcguirk et al., 1987; McGuirk and Ulsh, 1990; Knippertz and Martin, 2005) and cold air surges from the Mediterranean (e.g., Vizy and Cook, 2009, 2014), can account for a considerable amount of rainfall events, particularly in northern tropical Africa. Such interactions may even

arise in connection with AEWs, particularly during the late summer and early autumn seasons (Knippertz et al., 2003).

Mcguirk et al. (1987) define tropical plumes as "a continuous cloud band of upper or upper and middle clouds, at least 2000 km in length, crossing 15° N" in the northern hemisphere. Additionally, most studies agree that these cloud bands tend to be tilted in a southwest-northeast direction (Knippertz and Martin, 2005). Also known as 'moisture surges' or 'cirrus surges', tropical plumes are mostly found to be active between November and April, with the highest frequency in December (Knippertz, 2007). The orientation of these cloud formations is commonly employed in determining their source. During the winter months in the Northern Hemisphere, tropical plumes with a greater north-south orientation are more commonly observed, whereas those with a greater east-west orientation are favoured more in the summer months when the Intertropical Convergence Zone is more likely to shift towards the north. They often have a lifespan of about 1.5 to 2 days (Fröhlich et al., 2013).

While the definition provided by Mcguirk et al. (1987) is rooted in Pacific-based observations, Kiladis and Weickmann (1997) recognized the presence of tropical plumes in West Africa following the passage of a Rossby wave. During the transitional months of March and April, tropical plumes may significantly affect rainfall variability in northern subtropical-tropical Africa (Knippertz, 2003).

Cold air surges from the Mediterranean into northern Africa are a common source of rainfall variability in the extended summer months. Vizy and Cook (2014) define a cold air surge as "a shallow dome of relatively cold, dry air that advances equatorward, accompanied by a ridge of high pressure and strong meridional low-level winds". They often remain stationary over North Africa for about 10 days before propagating towards the equator at about 5.5 $ms^{-1}$. As they are usually colder than their surroundings by 1–6 K, they can significantly cool the environment by acting as a major energy sink in the tropics (Vizy and Cook, 2009; Garreaud, 2001). Such events have been documented to impact synoptic-scale rainfall over tropical Africa. Since the onset of the West African Monsoon is influenced by the intensity of the Saharan heat low, a powerful cold air surge penetrating deep into the Sahel during late Spring can weaken the heat low substantially and delay the onset of the WAM (Roehrig et al., 2011). While cold air surges in the summer are frequently preceded by heightened convective activity over the western Sahel, they lead to an approximately six-day-long monsoon break over the eastern Sahel. (Vizy and Cook, 2009).

#### 2.1.4.3 Madden-Julian Oscillation

The Madden-Julian oscillation (MJO; Madden and Julian, 1971, 1972) is the dominant intraseasonal mode of rainfall variability within the tropics. It was first discovered by Roland A. Madden and Paul R. Julian as a 40–50 day oscillation in the zonal wind in the tropical Pacific (Madden and Julian, 1971). Later studies by Zhang (2005)

discovered that this planetary-scale disturbance circumnavigates the tropics in about 30-90 days at approximately 5 $ms^{-1}$.

The influence of the Madden-Julian Oscillation (MJO) on rainfall patterns in tropical Africa has received widespread attention since the research conducted by Matthews (2000, 2004). The study put forth the idea that the disturbances generated by the MJO, namely eastward propagating Kelvin waves and westward propagating equatorial Rossby waves, complete a cycle around the globe and converge over Africa after approximately 20 days, leading to a peak in deep convection over Africa. Later studies (e.g., Gu, 2009; Pohl et al., 2009) found that the modulation of rainfall over the Guinea coast is more pronounced than that over the Sahel. In addition to enhancing rainfall, the MJO can also suppress convection. For instance, the study by Pohl et al. (2009) also discovered that the amplitude of MJO wave activity over the Sahel region is greater during its dry phase compared to its corresponding wet phases.

Although the MJO is a significant intraseasonal disturbance, numerous studies suggest that it also has the capability to affect rainfall variability on smaller timescales across Africa. For instance, experiments conducted with a global climate model (Lavender and Matthews, 2009) demonstrated that the MJO generates faster propagating disturbances like Kelvin and Equatorial Rossby waves that effectively boost monsoon convection over tropical Africa, confirming the conclusions of Matthews (2004). They also discovered that enhanced shear, induced on the African easterly jet by the MJO, increases AEW (see Section 2.1.4.1) activity, which modulates precipitation over tropical Africa, thereby indirectly impacting rainfall on shorter timescales.

#### 2.1.4.4 Tropical waves

Fluids on a rotating sphere, like the Earth, can develop planetary-scale disturbances that are generally limited to tropical or equatorial regions. These disturbances are formally referred to as tropical (or equatorial) waves (TWs). TWs can couple with convection (convectively-coupled TWs) and affect and regulate deep convection in the Earth's tropical atmosphere, leading to precipitation at different spatio-temporal scales (Wheeler and Kiladis, 1999; Kiladis et al., 2009; Knippertz et al., 2022, ; see next section for the theory). Since the spatio-temporal scales of different waves are different, their ability to modulate convection also varies. This dissertation considers both the theoretically described equatorial waves (Matsuno, 1966) and the additional large-scale tropical variabilities like the MJO and AEWs (see sections, 2.1.4.3 and 2.1.4.1, respectively) that show up as prominent modes in the wavenumber-frequency power-spectra (see Fig. **??**; Wheeler and Kiladis, 1999) as tropical waves.

Kelvin waves have been identified as a significant factor influencing rainfall variability at the synoptic scale in tropical Africa, particularly in West and Central Africa (Mekonnen and Thorncroft, 2016; Sinclaire et al., 2015). For instance, Mekonnen and Thorncroft (2016) present evidence that Kelvin waves account for up to 15% of daily rainfall

fluctuation in West Africa. This is supported by Ayesiga et al. (2021)'s recent study, which estimated modulation of approximately 5 mm per day. Furthermore, Lawton et al. (2022) documents the potential of Kelvin waves to trigger AEWs in the Sahel region, which are well known to initiate and modulate MCSs. Mekonnen et al. (2020) argue that, while African Easterly Waves have greater impacts on rainfall during wet years, Kelvin waves can enhance rainfall during dry years. They also note that dry years experienced more frequent and stronger Kelvin wave activity over Central Africa in contrast to wet years. A study conducted by Laing et al. (2011) in Central Africa found that, during their dry phase, Kelvin waves hinder the propagation of MCSs. However, during the passage of Kelvin waves in the wet phase, the MCSs propagate within their wave envelopes. It has also been observed that Kelvin waves have an effect on convective activity in Central Africa during the boreal spring season, in combination with enhanced MJO activity (Nguyen and Duvel, 2008; Laing et al., 2011).

Like the MJO, Equatorial Rossby (ER) waves impact the intraseasonal rainfall variability in tropical Africa. Janicot et al. (2010) identified two modes of ER waves which affect the West African summer monsoon during intraseasonal timescales. The first mode falls within a periodicity band of around 30-100 days, while the second mode lies within a band from 10-30 days. The former mode was found to be related to the activity of MJO over tropical Africa (Matthews, 2000, 2004). The latter mode propagates at a speed of approximately 6 $ms^{-1}$, exhibiting a period of approximately 15 days, and is frequently observed across the Sahel. However, it is not clearly understood whether these waves correspond to either the first or the second meridional mode or a combination of both as suggested by Yang et al. (2003). Not many studies exist that demonstrate the interaction between ER waves and AEWs or any other wave mode that can impact rainfall in tropical Africa at daily-subdaily timescales, unlike the clear evidence of MJO-AEW coupling (Matthews, 2004) and the possible connection between Rossby waves from the southern hemisphere and AEWs in the Sahel (Yang et al., 2018). The study by Schlueter et al. (2019a) describes an empirical approach to estimate the coupling between two tropical waves. It finds that low-frequency waves, such as the MJO and ER waves, can amplify high-frequency waves, such as tropical depression (TD) and MRG waves, in wet phases and suppress them during dry phases. Furthermore, a recent study conducted by Peyrillé et al. (2023) through a composite analysis suggests that there may exist a mechanism by which AEWs are triggered within ER envelopes. It should be noted, however, that the study in question solely focuses on extreme precipitation events, which significantly reduces the sample size.

The impact of mixed Rossby-gravity (MRG) waves on synoptic-rainfall variability remains unclear, except for one documented instance of a hybrid entity possessing characteristics of both the MRG waves and AEWs (MRG-AEW; Cheng et al., 2019), and a distinct case of interaction between MRG waves and AEWs in relation to Rossby waves originating from the southern hemisphere (Yang et al., 2018). The influence of eastward and westward propagating intertio-gravity waves (EIGs and WIGs, respectively) on African

rainfall is also unclear, mainly because of their high phase speeds. Nevertheless, it is now known that WIGs and squall lines (which are the main contributors to rainfall in the Sahel) are often indistinguishable (Tulich and Kiladis, 2012). This leads to the possibility of interpreting convectively coupled WIGs as squall lines in tropical Africa (Schlüter, 2019). As previously stated in section 2.1.4.1, Jung and Knippertz (2023) also suggests that WIGs could potentially be interpreted as high-speed AEWs.

## 2.2 Theory and identification of tropical waves

As highlighted by numerous studies (e.g., Roundy, 2012; Dias et al., 2018; Maier-Gerber et al., 2021), it is paramount to incorporate the generation and propagation of TWs into dynamical and statistical weather prediction methods to enhance tropical forecasting.

The theory of tropical waves relies on the *shallow-water* equations (SWEs) originating from the horizontal component of the linearization[1] of the primitive equations (the continuity equation, the equations regarding the conservation of momentum and a thermal energy equation), simplified for a basic state with no vertical shear (e.g., Matsuno, 1966; Wheeler and Kiladis, 1999). The SWEs govern motion in a three-dimensional, stratified atmosphere on a $\beta$-plane. The equations of motion and mass conservation in such a setup are as follows:

$$\begin{aligned} \frac{\partial u}{\partial t} - fv + g\frac{\partial h}{\partial x} &= 0, \\ \frac{\partial v}{\partial t} + fu + g\frac{\partial h}{\partial y} &= 0, \\ \frac{\partial h}{\partial t} + h_\mathrm{e}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) &= 0, \end{aligned} \tag{2.1}$$

where $u$ and $v$ are velocities in the positive $x$ and positive $y$ directions respectively and $h$ is the infinitesimal deviation in the elevation ($h_\mathrm{e}$) of the boundary layer. Note that $h_\mathrm{e}$ is more commonly referred to as the *equivalent depth* for the purpose of modelling motions of the atmosphere, and it will be referred to as such henceforth. However, it needs to be stated that $h_\mathrm{e}$ is in no way a physical depth in the real atmosphere, and it should only be considered as a parameter of the theory (Kiladis et al., 2009). $f$ is defined as the Coriolis parameter, and $g$ is the acceleration due to gravity. $f$ is assumed to be linearly proportional to the latitude ($y$) and hence:

$$f = \beta y, \tag{2.2}$$

where $\beta$ is the proportionality constant called the Rossby parameter. The Rossby parameter is defined as follows:

$$\beta = \frac{\partial f}{\partial y}, \tag{2.3}$$

[1]The linearization of the primitive equations yields Laplace's tidal equations

Since it is assumed that the atmosphere is in hydrostatic equilibrium, the equations in 2.1 can be rewritten as:

$$\begin{aligned} &\frac{\partial u}{\partial t} - \beta y v + g\frac{\partial \phi}{\partial x} = 0 \\ &\frac{\partial v}{\partial t} + \beta y u + g\frac{\partial \phi}{\partial y} = 0 \\ &\frac{\partial \phi}{\partial t} + gh_{\mathrm{e}}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0, \end{aligned} \tag{2.4}$$

where $\phi$ is the geopotential.

Since we are interested in zonally propagating wave solutions, $u$, $v$, and $\phi$ can be assumed to be sinusoidal in form as follows:

$$U = \hat{U} exp[i(kx - \omega t)], \tag{2.5}$$

where $U = (u, v, \phi)$ and $\hat{U} = (\hat{u}, \hat{v}, \hat{\phi})$ represent the deviations from $U$. In this context, the variables $k$ and $\omega$ are defined as the zonal wavenumber and (temporal) frequency, respectively. Using 2.5 to solve the equations in 2.4 yields a second-order differential equation in $\hat{v}$ only:

$$\frac{\mathrm{d}^2\hat{v}}{\mathrm{d}t^2} + \left(\frac{\omega^2}{gh_e} - k^2 - \frac{k}{\omega}\beta - \frac{\beta^2 y^2}{gh_e}\right)\hat{v} = 0, \tag{2.6}$$

for which the solutions of the form $v(y)$=$Ce^{\frac{-y^2}{2}}H_n(y)$, where $C$ is a constant, exist if the constant part inside the parenthesis of equation 2.6 satisfies:

$$\frac{\sqrt{gh_e}}{\beta}\left(\frac{\omega^2}{gh_e} - k^2 - \frac{k}{\omega}\beta\right) = 2n + 1 \qquad n = 0, 1, 2, \ldots, \tag{2.7}$$

where, $H_n(y)$ is the Hermite polynomial of order $n$.

Solving the SWEs leads to tropical wave modes characterized by four parameters: the meridional mode number $n$, the frequency $\omega$, the planetary zonal wavenumber $k$, and the equivalent depth $h_{\mathrm{e}}$. The theoretical dispersion relations derived from the solutions of the SWEs with any two parameters among $\omega$, $k$, and $h_{\mathrm{e}}$ allows the characterization of waves for a given wavetype and meridional mode, $n$ (Wheeler and Kiladis, 1999). The resulting solutions of SWEs can be symmetric or anti-symmetric about the equator depending on the field and $n$. For example, in fields such as precipitation and zonal wind, the solutions with odd $n$ are symmetric (Kelvin wave, n=1 equatorial Rossby wave, etc.), and the solutions with even $n$ are anti-symmetric (mixed Rossby–gravity

wave, n=0 eastward inertio gravity wave, etc.) while, in fields such as the meridional wind, these relationships are reversed (Tindall et al., 2006).

If we assume $c_e = \sqrt{gh_e}$, where $c_e$ is the phase speed of a non-dispersive wave, the solutions that satisfy Eq. 2.7 have the following dispersion relations:

1. Mixed Rossby-gravity (MRG) wave and $n = 0$ eastward-propagating intertio-gravity ($n = 0$ EIG) wave :

$$\begin{aligned} \omega_{\text{MRG}} &= c_e k \left( \frac{1}{2} + \frac{1}{2} \left( 1 + \frac{4\beta}{c_e k^2} \right)^{\frac{1}{2}} \right) \qquad k < 0 \\ \omega_{\text{EIG}} &= c_e k \left( \frac{1}{2} - \frac{1}{2} \left( 1 + \frac{4\beta}{c_e k^2} \right)^{\frac{1}{2}} \right) \qquad k \geq 0, \end{aligned} \tag{2.8}$$

   They are both solutions obtained by setting $n = 0$ in Eq. 2.7. MRG wave is a westward propagating mode and hence is valid only for negative values of k, while as the name suggests, EIG wave is eastward propagating.

2. $n = 1$ Equatorial Rossby (ER) wave

$$\omega_{\text{ER}} = \frac{-k\beta}{k^2 + \frac{3\beta}{c_e}} \qquad k > 0, \tag{2.9}$$

   ER waves are always westward propagating, hence the negative sign in the numerator. ER waves are the second slowest wave mode considered in this study after MJO. Higher meridional modes (i.e., $n > 1$) are also considered in some literature (e.g., Yang et al., 2003). However, they are not relevant to this study.

3. Kelvin wave

$$\omega_{\text{Kelvin}} = c_e k \qquad k > 0, \tag{2.10}$$

   Kelvin waves are eastward propagating wave modes. Since they require proximity to a lateral boundary, they are physically bound to regions very close to the equator[2], where the difference in the direction of the Coriolis force in the two hemispheres causes the equator to act as a rigid boundary. The solution for the Kelvin waves is obtained by substituting $v = 0$ in Eq. 2.4. However, it also emerges as a special solution in the solutions described by Matsuno (1966) by setting $n = -1$.

4. $n = 1$ Inertio-gravity (IG1) waves

$$\omega_{\text{IG}} = \pm\sqrt{3c_e\beta + c_e k^2}, \tag{2.11}$$

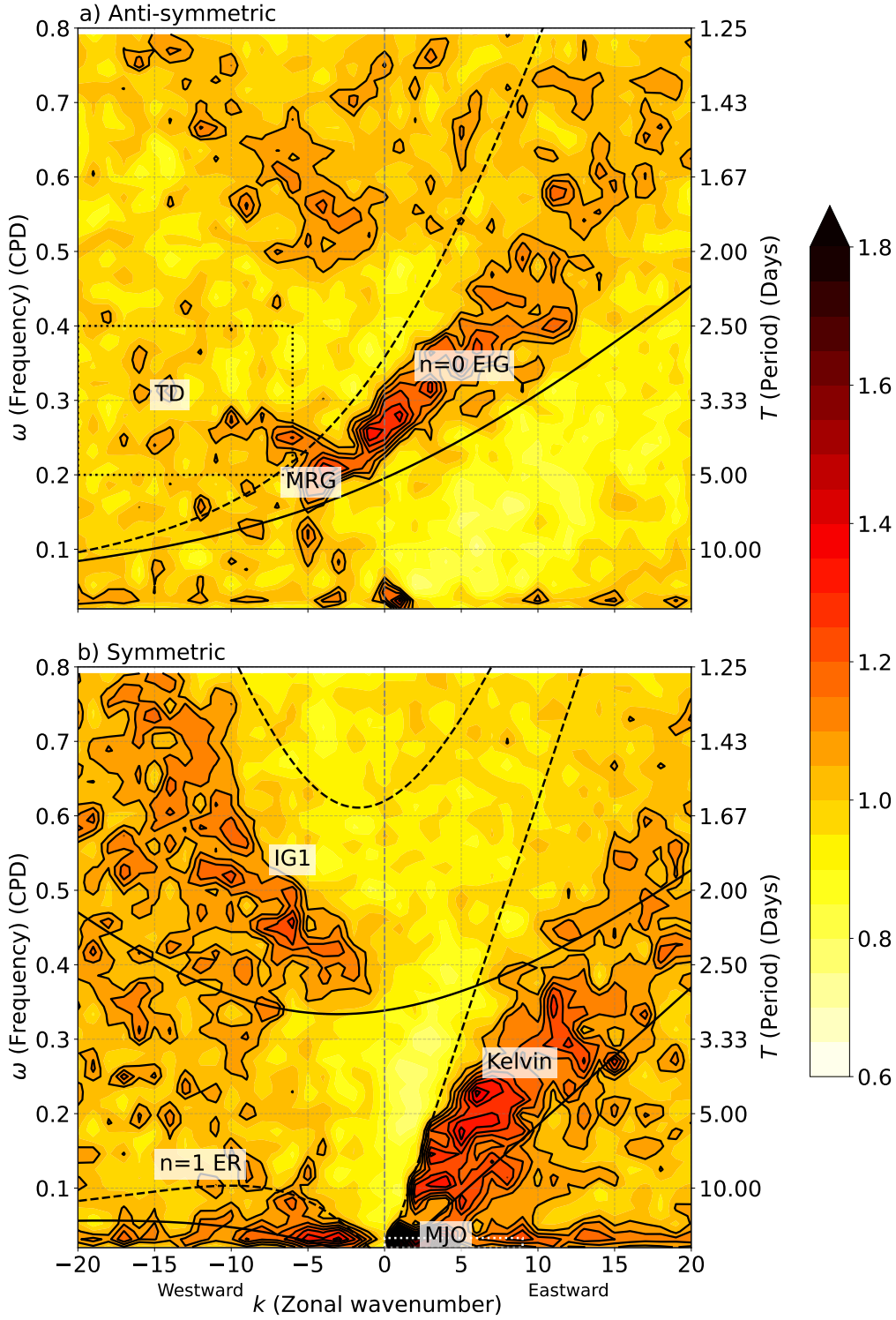
[2]We will not consider Kelvin waves that propagate along the coastlines in this study, since they are not a tropical wave

Inertio-gravity waves have an eastward (EIG1) and a westward (WIG1) component. Since only the WIG1 are known to affect $1-2$ day scale rainfall, in this thesis, we will only consider them (hereafter referred to as IG1).

While the theoretical solutions of the SWEs provide the dispersion relations for TWs such as the Kelvin wave, $n = 1$ equatorial Rossby (ER) wave, mixed Rossby-gravity (MRG) wave, $n = 0$ eastward inertio-gravity (hereafter, EIG) wave and the higher modes of intertio-gravity waves, power spectral analyses of satellite-observed fields such as the outgoing longwave radiation (hereafter, OLR), the narrow-band brightness temperature (hereafter $T_b$) which are considered as proxies for deep tropical convection (Wheeler and Kiladis, 1999; Kiladis et al., 2009), and precipitation have revealed the presence of the Madden Julian oscillation (MJO), as a prominent feature. The power spectra also show significant power for westward-moving 2.5-5 days TD type waves (Wheeler and Kiladis, 1999; Kiladis et al., 2009; Takayabu and Nitta, 1993). Although not a solution to the SWEs, the MJO (see section 2.1.4.3) is a significant component of tropical variability at the synoptic-meso scales, and TDs act as a close approximation to AEWs that modulate synoptic-meso scale precipitation over tropical Africa during the WAM (Schlueter et al., 2019a,b). The wavenumber-frequency power spectrum of rainfall field based on Global Precipitation Measurement Integrated Multi-satellitE Retrievals (GPM-IMERG) with the dispersion lines (based on the solutions) to identify tropical waves is shown in Fig. **??**.

## 2.3 Rainfall forecasting in tropical Africa

Rainfed agriculture is a significant source of food and income for livelihoods in tropical Africa, a region that would benefit greatly from accurate forecasts of rainfall on timescales from hours to seasons. Early warning systems for floods, droughts, and reservoir management are some other areas that could take advantage of improved forecasts. However, while economies in mid-latitudes have greatly profited from the 'quiet revolution' in numerical weather prediction (Bauer et al., 2015), such a development has not taken place in tropical Africa. In addition to the poor availability of in-situ observations (Fink et al., 2011; Parker et al., 2008), poor forecast skill is also due to lower intrinsic predictability in the tropics at meso- and synoptic scales (Judt, 2020), i.e., forecast errors at these scales grow faster than in mid-latitudes. This faster error growth relates to the convective nature of rainfall in contrast to frontal-dominated precipitation in mid-latitudes. Haiden et al. (2012) show that the skill of a six-day lead-time precipitation deterministic forecast in the extratropics is comparable to a one-day lead-time in the tropics. Specifically for tropical Africa, Vogel et al. (2018, 2020) and **?** demonstrate that – particularly in regions where long-lived Mesoscale Convective Systems (MCSs) provide the bulk of the annual rainfall (Mathon et al., 2002; Feng et al., 2021) – an Extended Probabilistic Climatology (EPC) based on records of past rainfall can outperform or reach skill comparable to state-of-the-art global Numerical Weather Prediction (NWP) models for 24-hour and 5-day ensemble rainfall forecasts.


a) Anti-symmetric
b) Symmetric
ω (Frequency) (CPD)
T (Period) (Days)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
1.25
1.43
1.67
2.00
2.50
3.33
5.00
10.00
TD
MRG
n=0 EIG
IG1
Kelvin
n=1 ER
MJO
1.8
1.6
1.4
1.2
1.0
0.8
0.6
−20
−15
−10
−5
0
5
10
15
20
Westward
Eastward
k (Zonal wavenumber)

While statistical postprocessing of NWP forecasts can correct for biases and dispersion errors, it still does not enhance the skill beyond that of EPC-based forecasts (Vogel et al., 2018). Compared to other parts of the tropics, the vast lowlands of West and Central Equatorial Africa stand out as the continental region with the lowest predictive skill of rainfall.

In addition to the low intrinsic predictability, deficiencies in operational systems may contribute to the low skill of rainfall predictions over tropical Africa. Current global NWP models at spatial resolutions of a few tens of kilometres parameterize moist convection. This leads to fundamental misrepresentations of MCSs and thus rainfall intensity and frequency over West Africa (Marsham et al., 2011; Pante and Knippertz, 2019), although recent improvements in parameterization may have alleviated this problem (Becker et al., 2021). Moreover, it is challenging to forecast the initiation of convection, as triggers tend to stem from unresolved small-scale processes. However, improved predictability may be achieved via specific conditions. For instance, Taylor et al. (2012) investigate the potential of soil moisture gradients as a storm-triggering mechanism, which gives rise to better short-term forecasts of convective initiation and tracks (Taylor et al., 2022). Studies such as Söhne et al. (2008) and Davis et al. (2013) have suggested that NWP forecasts may possess a higher level of predictive skill during regimes of strong synoptic-scale forcing, such as during the passage of tropical waves. As yet another example, Lafore et al. (2017) discuss the enhanced skill of precipitation forecasts related to orographic forcing over tropical Africa, as highlighted in Vogel et al. (2020) for the East African Highlands. However, northern tropical Africa and the Congo Basin have few orographic features. Thus, the overall inability of the forecast models to predict triggers and simulate the organization of convection are major reasons that NWP models struggle to produce skilful precipitation forecasts over West Africa (Fink et al., 2011). The poor model representation of convective processes also contributes to the fact that additional upper-air observations and, thus, an improved definition of the initial state are basically lost after 24 hours of model integration, as data denial experiments from the African Monsoon Multidisciplinary Analysis (AMMA) and Dynamics-Aerosol-Chemistry-Cloud Interactions in West Africa (DACCIWA) campaigns show (Agustí-Panareda et al., 2010; van der Linden et al., 2020).

The new generation of convection-permitting (CP) models allows a better representation of MCSs, the diurnal cycle of precipitation, and monsoon circulation over West Africa (Marsham et al., 2011; Pante and Knippertz, 2019). The latter authors note that the explicit simulation of Sahelian thunderstorms improves not only tropical, but also mid-latitude weather forecasts, at lead times of 5 to 8 days. Operational short-range forecasts with pan-Africa CP models were recently started at the UK Met Office, showing improved skill but also revealing problems with spin-up in the first 24 hours of integration (Hanley et al., 2021; Warner et al., 2023). Cafaro et al. (2021) demonstrate that CP ensemble forecasts for tropical East Africa show improved skill over deterministic forecasts at the 24-hour lead time but are under-dispersed in terms of the locations

of heavy rainfall and domain-average rainfall. Thus, while CP models improve the climatological representation of rainfall systems in Africa, the gain in predictive skill appears limited, and the computational costs are enormous (Senior et al., 2021).

Given the general challenge of forecasting non-topographically triggered mesoscale convection, it is conceivable that windows of enhanced predictability may exist when large-scale, more predictable trigger and forcing mechanisms act. For boreal summer, Kniffka et al. (2020) note enhanced model skill in rainfall predictions over southern West Africa during the DACCIWA campaign when long-lived synoptic vortices were present. It is also well-known that tropical waves, in particular AEWs, modulate rainfall in West Africa during the boreal summer wet season (Reed et al., 1977; Fink and Reiner, 2003; Lavaysse et al., 2006; Schlueter et al., 2019a,b). These and other studies suggest that the waves do not directly trigger convection but change the larger scale (thermo-)dynamical environment and, thereby, enhance the odds of triggering and organisation of MCSs.

The links between waves and rainfall in West Africa suggest two alternative avenues of forecasting precipitation. Firstly, one can leverage the more predictable large-scale waves in mixed statistical-dynamical forecast approaches, similar to methods developed for tropical cyclone occurrence in the North Atlantic Ocean by Maier-Gerber et al. (2021), or in wave pattern-based calibration approaches (de Andrade et al., 2021). A second approach is to use purely data-driven forecast models. Such models have a long tradition going back to the times prior to NWP, as they are computationally cheap and tend to be well-calibrated by construction. Early approaches often relied on Markov processes, as in Gabriel and Neumann (1962) and Gates and Tong (1976). More recently, advanced statistical and machine learning tools have been employed (Diez-Sierra and del Jesus, 2020; Ravuri et al., 2021).

While studies of this type may not seek physical explanation, their successes rely on the spatio-temporal coherence of rainfall and systematic relationships to other meteorological variables. Following this avenue, Vogel et al. (2021) show that a relatively simple, purely data-driven logistic regression model for 24-hour precipitation occurrence, which draws on spatio-temporal patterns of daily rainfall, outperforms forecasts based on climatology or NWP models during the summer monsoon season in West Africa. The model relies on the identification of grid points with the strongest correlation to a target grid point at lags of one and two days. The success of this method suggests that statistical models can reliably represent the systematic relationships between rainfall and circulation features that dynamical models appear to struggle with (Elless and Torn, 2018). Vogel et al. (2021) argue that their approach can, in principle, be employed whenever there is a dominant coherent forcing that triggers and modulates rainfall.

Yet another option is to utilise tropical waves as predictors in statistical models since they are a significant source of predictability within the tropics Kiladis et al. (2009). A recent study by Li and Stechmann (2020) utilized tropical waves, which were modelled as damped oscillators, to predict tropical rainfall. The authors contend that employing

tropical waves as predictors might extend the forecast lead time by up to two times compared to using only non-wave predictors. Nonetheless, it remains unclear whether such assertions are valid for tropical Africa, considering the multiple reasons outlined in this section. Since many studies (e.g., Wheeler and Weickmann, 2001; Janiga et al., 2018) present techniques to forecast tropical waves accurately, it is logical to conclude that such predictability can be leveraged to forecast rainfall over tropical Africa utilising statistical models.

# 3. Research questions

From the previous chapter, it is understood that rainfall prediction in northern tropical Africa, especially in the semi-arid Sahel and the Congo Basin, is very challenging. Mainly consisting of flat terrain without the requisite orographic features to induce deep convection (Lafore et al., 2017), the convective processes in this region are primarily driven by small-scale surface processes. The small-scale convective cells that are formed due to such processes develop into mesoscale convective systems when the necessary (thermo-)dynamic conditions are present (Mathon et al., 2002; Fink et al., 2006; Vogel et al., 2018). Despite decades of technological advancements in computing, the present-day NWP models are incapable of resolving such processes and cannot provide a reliable representation of the genesis of convective events over tropical Africa (Fink et al., 2011). While ensemble prediction systems have somewhat alleviated this issue, the fundamental problem remains unsolved. Even the ECMWF precipitation ensemble forecasts, which surpass their competitors in tropical Africa (Wang et al., 2023), are unable to outperform simple climatology-based forecasts. Statistical post-processing techniques for correcting errors and biases also prove ineffective in achieving the required level of skill, particularly in regions with precipitation dominated by MCSs (Vogel et al., 2018, 2020; **?**). The convection-permitting (CP) ensemble forecasts provide an improved representation of MCSs but lack robustness in predicting rainfall events accurately at the correct locations (Cafaro et al., 2021). Moreover, the enormous computational expenses associated with operating CP forecasts with numerous ensemble members outweigh the marginal gain in predictive skill (Senior et al., 2021).

A potential solution to these challenges is to utilise solely data-driven techniques, specifically those that rely only on observational data. Whilst previously developed using Markov processes well before the widespread use of NWP models (Gabriel and Neumann, 1962; Gates and Tong, 1976), purely data-driven models have not been as commonly used in weather forecasting as their NWP equivalents, partially due to their lack of emphasis on determining physical causality for their results. The effectiveness of such models is rather heavily reliant on the existence of coherent relationships between rainfall and other meteorological variables. With recent technological advancements, statistics-based and machine-learning tools have become increasingly accessible to the scientific community, leading to an upsurge in the popularity of data-driven approaches. Studies such as Diez-Sierra and del Jesus (2020) and Ravuri et al. (2021) reinforce this

viewpoint. A recent study by Vogel et al. (2021) demonstrates that a straightforward, purely data-driven, logistic regression-based model for predicting the likelihood of daily rainfall in northern tropical Africa is able to outperform ensemble forecasts based on NWP models during the summer monsoon season. The model in question can draw coherent rainfall propagation patterns over this region from the last two days' rainfall information, which mainly accounts for the improved predictive skill.

Motivated by such examples, this thesis aims to furnish meteorological evidence to support the enhanced skill of statistical forecasting, particularly with respect to short-term precipitation over northern tropical Africa. The chapters 6 and 8 comprise thorough depictions of the findings using the statistical forecasts produced via the models described in Chapter 5, supported where possible by meteorological explanations.

Chapter 6 builds on the foundation established by Vogel et al. (2021) and aims to predict daily rainfall occurrence in tropical Africa using a logistic regression (Logistic) model trained on rainfall data from the previous three days. By comparing the skill of the Logistic forecast to the benchmarks, including the forecast based on climatology, the raw NWP-based forecast, and a statistically postprocessed NWP-based forecast, the chapter aims to address the following research questions:

**RQ1a** How does the skill of the Logistic forecast compare to that of the climatology and the NWP-based forecast (both raw and postprocessed)?

**RQ1b** How much predictive value do lags of one, two and three days provide?

**RQ1c** How can the performance of the Logistic forecast be explained using coherent propagation of rainfall patterns and meteorological variables like Convective Available Potential Energy (CAPE), shear and Convective Instability (CIN)?

As discussed in the previous chapter, studies such as Söhne et al. (2008) and Davis et al. (2013) have suggested that NWP forecasts may possess a higher level of predictive skill during regimes of strong synoptic-scale forcing, which may be in part attributed to the influence of tropical waves, particularly their phases. Therefore, the phases of tropical waves (i.e., whether wet or dry) are a crucial factor for effectively utilising predictors based on tropical waves in NWP models. This has been discussed at length in the previous chapter. When implementing models that rely on tropical waves to provide valuable predictive information, it is justified to assume that the wave-phases play prominent roles in modifying the (thermo-)dynamical conditions over tropical Africa in order to trigger and potentially modulate the amount of precipitation. This assumption aligns with the findings of Schlueter et al. (2019a,b).

Motivated by these prospects, Chapter 7 aims to answer the following research questions:

**RQ2a** How do the predictive abilities of various tropical waves vary across different regions of tropical Africa? Are the results consistent with the literature?

**RQ2b** To what extent do the phases of tropical waves impact the skill of predicting precipitation occurrence and amounts?

Chapter 6 focused on a distinct objective: to produce a statistical model for predicting daily rainfall occurrence in tropical Africa, competitive with the benchmarks, supported with meteorological reasoning. However, a more extensive and ultimately more beneficial forecasting method would instead predict daily rainfall accumulation in the region. The findings of Chapter 7 demonstrate that the phases of convectively coupled tropical waves modulate the skill of NWP model-based forecasts of daily rainfall over tropical Africa. It is reasonable to assume that the systematic propagation of precipitation patterns caused by tropical waves can provide satisfactory predictive ability, which statistical models can utilise to produce accurate forecasts of daily rainfall amounts in tropical Africa.

Chapter 8 develops statistical models, using exclusively the predictive information from seven convectively coupled tropical waves, specifically MJO, TDs, Kelvin, ER, MRG, EIG, and IG1 waves. These models predict daily rainfall accumulations in tropical Africa and consist of a gamma regression model and a convolutional neural network (CNN) model (see Chapter 5 for details on these models). Comparing the skill of these forecasts with benchmarks, such as two climatology-based forecasts, an NWP model-based ensemble forecast and an NWP model-based probabilistic forecast, this chapter aims to address the following research questions:

**RQ3a** How do the skills of gamma regression and the CNN models measure up against those of the forecasts based on climatology and the NWP model-based forecasts (both ensemble and probabilistic)?

**RQ3b** How does the skill of gamma regression forecast compare against that of CNN forecast?

By answering these research questions, this dissertation aims to exhibit the merits of employing statistical models trained on observational data for forecasting daily rainfall in tropical Africa, supported by meteorological reasoning.

# 4. Data and methods

## 4.1 Observational data and ECMWF ensemble forecasts

To represent precipitation, we use the satellite-based, gridded GPM IMERG final version product (Hou et al., 2014; Huffman et al., 2020) to represent precipitation. GPM IMERG uses both radar-calibrated microwave radiance from polar-orbiting satellites and infrared radiance from geostationary satellites to create grids with half-hourly temporal and 0.1° spatial resolution. We conservatively remap to $1° \times 1°$ spatial resolution in order to suppress small-scale noise in the precipitation field and ease the computational burden, and we consider daily accumulations (0600 UTC to 0600 UTC) that reflect the diurnal cycle of convection. To define precipitation occurrence in chapter 6, we use a binary cutoff at 0.2 mm of daily accumulation. To compute the TW-filtered data based on the precipitation field for use in chapters 8 and 7, we use the 6-hour mean precipitation data from 00UTC-06UTC.

The ERA5 dataset (Hersbach et al., 2020) is used for horizontal wind at 600 and 925 hPa, Convective Available Potential Energy (CAPE), and Convective INhibition (CIN). To match the spatial resolution of GPM-IMERG in this study, the ERA5 data have been remapped from the native $0.25° \times 0.25°$ resolution to the $1° \times 1°$ grid using bilinear interpolation.

We compare forecasts with our data-driven model to the operational global 24-h ensemble forecast from the ECMWF, which consists of 50 perturbed members and a control run at a spatial resolution of $0.25° \times 0.25°$ (ECMWF Directorate, 2012). We conservatively remap to $1° \times 1°$ resolution and daily accumulations from 0600 UTC to 0600 UTC, and we refer to this forecast as ENS. As Wang et al. (2023) demonstrate, ECMWF forecasts outperform National Centers for Environmental Prediction–Global Forecast System (NCEP–GFS) forecasts of precipitation over Africa.

In chapter 6, we restrict all datasets to a latitude–longitude band comprising 12° S to 30° N and 68° W to 75° E. The spatial domain considered thus covers most of tropical Africa, the adjacent oceans, as well as parts of South America and Southwestern Asia (Fig. 5.2). The use of a significantly larger zonal extent compared to the meridional extent is motivated by our interest in identifying zonally propagating features. Temporally, the observations cover the following time periods: summer monsoon season (July, August,

and September), pre-monsoon season (May and June) and two transition seasons (March–April and October–November) where applicable in the 19-year period from 2001 to 2019, and the forecasts cover the 13 years from 2007 to 2019, reflecting the ENS data availability in the THORPEX Interactive Grand Global Ensemble (TIGGE: Bougeault et al., 2010; Swinbank et al., 2016) archive. In chapters 8 and 7, we restrict the analysis to latitude–longitude band comprising 0° S to 18° N and 25° W to 35° E to represent northern tropical Africa, but use precipitation data over all longitudes in the tropical latitude band to filter tropical waves.

## 4.2 Tropical wave filtering

A variety of methods have been proposed to identify and filter out TWs (e.g., Wheeler and Kiladis, 1999; Yang et al., 2003; Roundy et al., 2009; Yang et al., 2003; Roundy et al., 2009). Knippertz et al. (2022) offer a comprehensive summary of six such approaches with their pros and cons. In this study, we have decided to utilise the technique detailed in Wheeler and Kiladis (1999) and later modified by Janiga et al. (2018), as it allows the filtering of non-theoretical wave modes such as the MJO and TD. In this method, the wavenumber-frequency power spectrum is computed using a 2D fast Fourier transform (FFT) applied to the raw field along time and longitude, as shown in Fig. **??**. To identify the filter windows for every TW, we superimpose the dispersion relations based on the solutions of SWEs for various TWs and determine the ranges of the time period, wavenumber, and $h_e$ required as shown in Table 4.1. By performing an inverse FFT and setting zeros to all wavenumbers and frequencies outside of the designated windows, we are able to obtain the real-valued field for each wave mode.

| Wave | Time period (days) | Wavenumber | Equivalent depth (m) |
|---|---|---|---|
| Kelvin | 2.5 – 20 | 1 – 14 | 8 – 90 |
| MJO | 30 – 96 | 0 – 9 | – |
| EIG | 1.82 – 5 | 0 – 14 | 8 – 90 |
| MRG | 3 – 8 | -10 – -1 | 8 – 90 |
| TD | 2.5 – 5 | -20 – -6 | – |
| IG1 | 1.4 – 2.5 | -20 – -1 | 8 – 90 |
| ER | 9 – 72 | -10 – -1 | 8 – 90 |

**Table 4.1:** The wavenumber, time period, and equivalent depth ranges utilized to filter various tropical waves, as per Janiga et al. (2018).

To generate predictors using the filtered wave modes, the following steps are taken: a) Before filtering, the data set is divided into training and test sets in a similar manner as described in section 5.2.1, since we are operating in cross-validation mode, b) After filtering, the first and last three years are removed from each training set to reduce the edge effect due to the FFT and inverse FFT procedures used. To serve the same purpose, each test set is padded with three years of zeroes on either side before filtering, and c) The local phases and amplitudes are computed as explained in the next section.

## 4.3 Statistical postprocessing

Statistical postprocessing is a crucial aspect of modern weather forecasting. It involves rectifying various biases and errors in forecasts to enhance their calibration. This is usually accomplished by utilizing statistical techniques to learn from the deviations between past forecasts and past observations by fitting a suitable statistical model and then applying these corrections to future forecasts.

Some commonly used statistical postprocessing methods used in the field of weather forecasting are Ensemble Model Output Statistics (EMOS; Gneiting et al., 2005) and Bayesian Model Averaging (BMA; Raftery et al., 2005). EMOS employs a distributional regression technique, assuming a parametric distribution for the target variable, based on the predictors (Schulz, 2023). Depending on the target variable, various distributions are employed. For instance, temperature is often modelled using a Gaussian distribution, whilst zero-censored distributions are used for variables like precipitation, which are never negative. The ensemble members, typically used as predictor variables, are fitted to the assumed distribution and their parameters are estimated by optimising some form of proper score, such as the continuous ranked probability score (CRPS; see Section 4.5.3.4). BMA, on the other hand, produces probabilistic forecasts from ensemble forecasts in the form of probability density functions (PDFs) (Sloughter et al., 2007). The BMA method models each member of the ensemble using a distinct distribution and then combines parameter estimates by weighting them according to the model probability.

In this thesis, we employ a recently introduced Isotonic Distributional Regression (IDR; Henzi et al., 2021) technique. Unlike the methods described above, IDR does not require the assumption of a distribution and hence requires no parameter tuning. Furthermore, in the case of probability forecasts for binary outcomes, IDR coincides with isotonic calibration (Zadrozny and Elkan, 2002), which is an established calibration technique in machine learning. Since this method permits the postprocessing of precipitation by solely employing probabilities of occurrence, it is computationally less demanding than alternative techniques while still maintaining robustness.

## 4.4 Uncertainty quantification via EasyUQ

Easy Uncertainty Quantification (EasyUQ; Walz et al., 2022) is a novel tool that transforms deterministic forecasts, such as those produced by statistical models or NWP, into a probabilistic distribution. No knowledge of the input parameters that generated the deterministic forecast is required beyond the forecast and observations. The basic version of EasyUQ uses IDR (Henzi et al., 2021) to generate discrete predictive distributions without the need for parameter tuning. The IDR technique includes the additional advantage of delivering readily calibrated results. The EasyUQ enables us to easily determine prediction intervals. In this thesis, we shall use the interval between the 5th and 95th percentiles to display the range of the predictive distribution.

Since we are operating in 'leave-one-out' cross-validation mode, we will employ the same strategy that was used earlier to train EasyUQ, namely: 1) Split the deterministic forecast and observation into train-test sets and train EasyUQ on data from all available years except the year for which the probabilistic forecast is issued. 2) We repeat this process for every train-test split and merge the outcomes to achieve a complete probabilistic forecast spanning the years 2007 to 2016. 3) Based on the recommendation of Walz et al. (2022), we will employ CRPS as a metric for assessing the accuracy of the newly produced probabilistic forecast. The use of CRPS is preferred for this purpose due to two reasons: Firstly, it is reported using the same unit as the forecast. Secondly, it reduces to mean absolute error (MAE) for single-valued outputs, thereby enabling direct comparison between the deterministic and probabilistic forecasts (Gneiting and Raftery, 2007; Walz et al., 2022).

## 4.5 Forecast evaluation metrics

We now describe the various evaluation metrics that we use in this dissertation to assess the skill of the forecasts (both statistical and dynamical).

### 4.5.1 Rainfall occurrence

We mainly employ two evaluation metrics to assess the quality of the forecasts of rainfall occurrence. The metrics are: 1) The Brier (skill) score and 2) Reliability diagrams.

#### 4.5.1.1 Brier (skill) score

Given a set of N forecast-observation pairs, where each forecast represents the likelihood of an event occurring (e.g. whether the level of rainfall will surpass a certain threshold), and each observation denotes a binary outcome (in the aforementioned example, 1 if the level of rainfall surpassed the threshold and 0 if it did not), the Brier score (BS; Brier, 1950) can be utilized to evaluate the calibration of the forecast. The BS is defined as:

$$BS = \frac{1}{N}\sum_{i}^{N}(y_i - \hat{y}_i)^2 \tag{4.1}$$

where $y_i$ represents the binary observation and $\hat{y}_i$ is the probability (between 0 and 1) of the forecast for the $i^{\text{th}}$ instance. The Brier Score is a metric ranging from 0 to 1, with lower values indicating better performance. A perfect forecast will have a BS of 0, while the worst forecast will have a BS of 1. A forecast based on random chance alone, with $\hat{y}_i$ equal to 0.5, will have a BS of 0.250. Therefore, forecasts with a BS greater than 0.250 are, strictly speaking, worse than random forecasts. This suggests that good forecasts should aim for a BS of less than 0.250.

The value of the BS can be broken down (Murphy, 1977; Sanders, 1963) into three components, specifically reliability (REL), resolution (RES), and uncertainty (UNC):

$$BS = REL - RES + UNC \tag{4.2}$$

where each term can be defined as:

$$\begin{aligned} REL &= \frac{1}{N} n_k \sum_{k}^{K} (f_k - y_k)^2 \\ RES &= \frac{1}{N} n_k \sum_{k}^{K} (y_k - \bar{y})^2 \\ UNC &= \bar{y}(1 - \bar{y}) \end{aligned} \tag{4.3}$$

where $K$ is the number of bins into which the forecast probabilities are divided, $n_k$ is the number of forecasts within the $k^{\text{th}}$ bin, $f_k$ and $y_k$ are the mean forecasts and observations respectively in the $k^{\text{th}}$ bin and $\bar{y}$ is the observed rate at which the event occurs in climatology.

$REL$ assesses how accurate the forecast probabilities are by measuring how close they are to the true probabilities and is negatively oriented, i.e. the lower the value, the better. A good example of a reliable forecast is where there's a 60% probability of rainfall exceeding a certain limit forecasted, and this happens 6 times out of 10.

$RES$ quantifies the deviation of the conditional probabilities of different forecasts from the climatology and is positively oriented, i.e. higher values are better.

$UNC$ measures the inherent uncertainty and does not depend on the forecasts.

However, as this method for computing the BS relies heavily on the number of bins chosen, Dimitriadis et al. (2021) has proposed a more robust approach that utilises the newly introduced optimal binning computed by the CORP method. This approach decomposes the BS into three components: miscalibration (MCB), discrimination (DSC) and, as earlier, uncertainty (UNC), where the miscalibration and discrimination terms are analogous to the traditional reliability and resolution terms, respectively. For improving the BS of a forecast, one should aim to decrease the MCB and increase the DSC.

While the Brier score (BS) is used to assess the performance of an individual forecast, it is frequently necessary to evaluate a forecast in relation to a reference forecast. It provides a measure of the relative accuracy of the forecast in question. In this respect, the Brier skill score (BSS) is introduced. The BSS is defined as:

$$BSS = \frac{BS_{\text{reference}} - BS}{BS_{\text{reference}} - BS_{\text{perfect}}} = \frac{BS_{\text{reference}} - BS}{BS_{\text{reference}} - 0} = 1 - \frac{BS}{BS_{\text{reference}}} \tag{4.4}$$

where $BS_{\text{reference}}$ represents the BS of the reference forecast and $BS_{\text{perfect}}$ represents the BS of a perfect forecast, which is zero. In contrast to BS, BSS is positively oriented and has a range between $-\infty$ and 1, with 1 being the maximum achievable BSS.

#### 4.5.1.2 Reliability diagrams

Calibration is a vital element that determines the credibility of a forecast. According to Gneiting et al. (2007), calibration is the "statistical consistency between the predictive

distributions and the associated observations". This can be understood by a straightforward example: Suppose a forecast predicts a 60% likelihood of rainfall at a particular location. If we calculate the likelihood of rainfall based on observational data (assuming a reasonable sample size) and estimate it to be approximately 60%, we may conclude that the forecast is 'well-calibrated'.

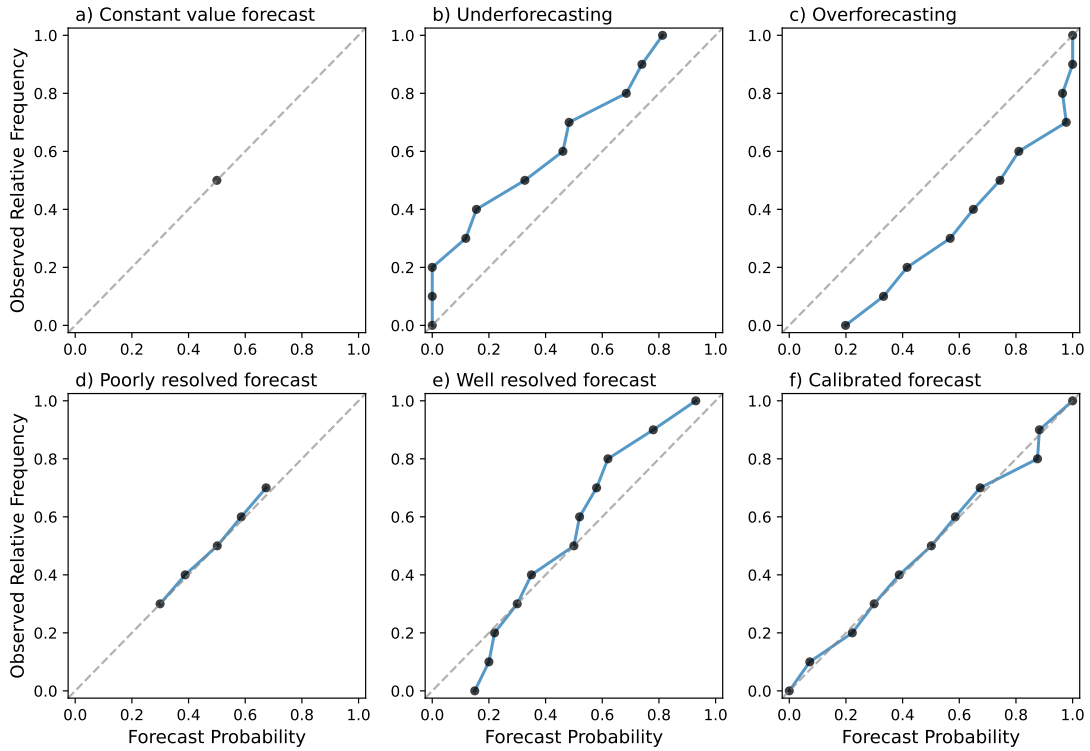



**Figure 4.1:** Illustrations of reliability diagrams for forecasts of binary events calibrated at different levels. Further information can be found in the text.

Reliability diagrams, also known as calibration curves, are utilised as a common technique to visualise the calibration of probabilistic forecasts of binary events (Wilks, 1990). This method involves plotting the projected probability issued by the forecasts (Forecast Probability) on the x-axis against the estimated Observed Relative Frequency (or Conditional Event Probability) on the y-axis. Such diagrams are illustrated in Fig. 4.1. The dashed line drawn along the diagonal is representative of a perfectly calibrated forecast. In other words, deviations from the diagonal illustrate the miscalibration. The reliability diagram displayed in Figure 4.1f serves as an example of a well-calibrated forecast where the observed relative frequency corresponds almost exactly to the forecast probability, with some deviations that are consistent with sampling variabilities (Wilks, 2011).

All additional reliability diagrams presented in Figure 4.1 display frequent forecasting issues. A model that provides a constant forecast will be represented as a single point on the diagram (Fig. 4.1a). If the model is based on climatology, the point will appear somewhere on the diagonal, as climatological forecasts are usually well-

calibrated, provided they were subjected to detrending. Figs. 4.1b–c are examples of forecasts with 'unconditional' biases. In Fig. 4.1b (Fig. 4.1c), the reliability curve (blue) is located entirely on the left (right) of the diagonal, implying that the forecasts issued by that model are always too small (large) relative to the observed relative frequencies, or the model is underforecasting (overforecasting). In the context of forecasting rainfall occurrence, the term dry bias (wet bias) is also sometimes used to describe underforecasting (overforecasting) (Wilks, 2011). Figures 4.1d–e exemplify 'conditional' biases, which are dependent on the forecasts themselves (Wilks, 2011). Specifically, Fig. 4.1d (Fig. 4.1e) illustrates a poorly-resolved (well-resolved) forecast, whereby the model is only able to forecast a subset of the entire probability distribution. Climatology-based predictions frequently exhibit such characteristics as much of the variability in the data is lost in the estimation of the climatology. On the contrary, a well-resolved forecast has the ability to predict all probabilities but is not properly calibrated. Additionally, such forecasts often demonstrate a lack of consistent shift, displaying both over-and-underforecasting tendencies.

Traditional reliability diagrams usually group forecast-observation pairs and average them into bins according to the forecast probability. However, this approach often creates a reliability diagram that is highly dependent on the choice of bins, which can lead to irreproducibility and unstable reliability diagrams (see Fig. 6.4 in chapter 6). To address this issue, using the pool-adjacent-violators (PAV) algorithm, Dimitriadis et al. (2021) developed the Consistency Optimality Reproducibility PAV (CORP) method that utilises optimally computed bins and assumes a monotonic, non-decreasing shape of conditional event probability.

### 4.5.2 Benchmarking: EPC15, ENS, and ENS-IDR forecasts (Rainfall occurrence)

In order to benchmark the performance of the purely data-driven Logistic probability of precipitation forecasts, we compare it to an Extended Probabilistic Climatology with a 15-day window (EPC15; **?**), to the raw ECMWF ensemble (ENS), and to a statistically post-processed ensemble (ENS-IDR). For a fair comparison in cross-validation mode, we treat the EPC15 forecast in the same way as the Logistic forecast, i.e., the probability of precipitation is determined from a ±15-day window around the target date, based on observations from 2001 to 2019 except the year for which the prediction is issued. This is repeated for each of the years from 2007 to 2019, for which we assess and compare the four types of forecasts.

The operational ECMWF ensemble consists of 50 perturbed members and a control run, so the ENS probability of precipitation forecast equals the number of ensemble members that predict a precipitation accumulation of at least 0.2 mm divided by 51. Statistical postprocessing can correct systematic deficiencies in NWP ensemble output (Vannitsem et al., 2018), and various methods, ranging from simple linear regression to sophisticated neural-network-based techniques, have been proposed for this purpose. In

this study, we use Isotonic Distributional Regression (IDR; Henzi et al., 2021) since, unlike every single competitor, IDR does not depend on user decisions and has been shown by Henzi et al. (2021) to be competitive with Bayesian Model Averaging (BMA; Sloughter et al., 2007) and Ensemble Model Output Statistics (EMOS; Scheuerer, 2014) for precipitation (Henzi et al., 2021). Furthermore, in the case of probability forecasts for a binary outcome, such as the occurrence of precipitation, IDR coincides with isotonic calibration (Zadrozny and Elkan, 2002), which is a widely used, classical method in machine learning. We retain the cross-validation setting used to generate the Logistic and EPC15 forecasts, i.e., we train IDR[1] on observations from the monsoon seasons (July–September, hereafter JAS) in 2007 to 2019 except the year for which the prediction is issued, and repeat for each of the evaluation seasons from 2007 to 2019.

To assess whether or not an observed difference in the mean Brier score of the Logistic forecast and a benchmark forecast could be explained by chance alone, even though the theoretical skill is equal, we employ the Diebold–Mariano test (Diebold and Mariano, 2002) in concert with the Benjamini–Hochberg correction (Benjamini and Hochberg, 1995), as recommended by Wilks (2016) for the spatio-temporal setting at hand.

### 4.5.3 Rainfall accumulation

We employ various evaluation metrics to assess the quality of the forecasts of rainfall accumulation. The metrics we employ to evaluate the deterministic forecasts produced include: 1) the Pearson correlation coefficient, 2) the Taylor score (Taylor, 2001), and 3) the Kling-Gupta efficiency (KGE) (Gupta et al., 2009; Knoben et al., 2019). We quantify the uncertainty of these deterministic forecasts in relation to the observations using EasyUQ (Walz et al., 2022) and generate probabilistic forecasts and assess their skill through the CRPS (Matheson and Winkler, 1976).

#### 4.5.3.1 Pearson correlation

To assess the relationship between the deterministic forecasts ($\hat{y}$) and the observations ($y$), we employ the Pearson correlation coefficient. The Pearson correlation coefficient ($\rho$) is defined as below:

$$\rho = \frac{\sum(y - \bar{y})(\hat{y} - \bar{\hat{y}})}{\sqrt{\sum(y - \bar{y})^2 \sum(\hat{y} - \bar{\hat{y}})^2}} \tag{4.5}$$

where $\bar{y}$ and $\bar{\hat{y}}$ are the mean of $y$ and $\hat{y}$ respectively.

#### 4.5.3.2 Taylor Score

Taylor (2001) defines the skill of a deterministic forecast as dependent on two factors: firstly, how closely the forecasted values match the observed (low root mean squared

[1]To implement IDR, we use the `isodisreg` package for Python; see https://github.com/evwalz/isodisreg.

error, RMSE) and secondly, how accurately the forecast represents observed patterns and trends (high correlation). It is noteworthy that since the RMSE is known to be biased in the presence of outliers, Taylor (2001) explicitly removes the bias when calculating the RMSE (also called centred RMSE) to address this. As such, they provide a score definition that combines these two aspects and varies from zero (least skilful) to one (most skilful). They define this as the Taylor Score $S$, and it is given as follows:

$$S = \frac{4(1+\rho)}{(\hat{\sigma} + \frac{1}{\hat{\sigma}})^2(1+\rho_0)} \tag{4.6}$$

where $\hat{\sigma}$ is defined as the forecast variance normalized with the observed variance, and $\rho_0$ is the maximum correlation possible. For our purpose, we assume that $\rho_0$=1. So as $\hat{\sigma} \to 1$ and $\rho \to 1$, $S \to 1$. On the other hand, when $\rho \to -1$ or $\hat{\sigma} \to \infty$, $S \to 0$.

#### 4.5.3.3 Kling-Gupta efficiency

Gupta et al. (2009) introduced the Kling-Gupta efficiency (KGE) metric as an alternative to traditional approaches to evaluating the skills of hydrological models. The KGE, similar to the Taylor score, is based on the premise that a competent forecast should have minimal errors and a strong correlation with observations. The KGE is calculated using the following formula:

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \tag{4.7}$$

where $\alpha = \frac{\sigma_{\text{forecast}}}{\sigma_{\text{observation}}}$ is called the flow variability error, $r$ is the Pearson correlation and $\beta^2 = \frac{\mu_{\text{forecast}}}{\mu_{\text{observation}}}$ is the bias term. Here $\sigma_{\text{forecast}}$ and $\sigma_{\text{observation}}$ are the standard deviation of the forecast and observation, respectively, and $\mu_{\text{forecast}}$ and $\mu_{\text{observation}}$ are the mean of the forecast and observation, respectively.

It should be pointed out, however, that different authors have different interpretations of the KGE (see Knoben et al. (2019) for details) as to what value should be taken as a measure of a good forecast. Nevertheless, they all concur that a higher value of the KGE relative to another forecast (e.g. from another model) is considered to be better, and a lower value is considered to be worse. Therefore, this dissertation will use KGE as a relative measure to compare various deterministic forecasts.

#### 4.5.3.4 Continuous ranked probability (skill) score

To assess the quality of the likelihood predictions of rainfall accumulations, the continuous ranked probability score (CRPS Gneiting and Raftery, 2007) prevails as the most widespread metric employed. In essence, CRPS can be implemented when the likelihood prediction manifests as a cumulative distribution function (CDF) and the observation is a scalar. Such evaluation of forecasts can be used to determine its efficacy. The

[2] not to be confused with the Rossby parameter defined in section 2.2

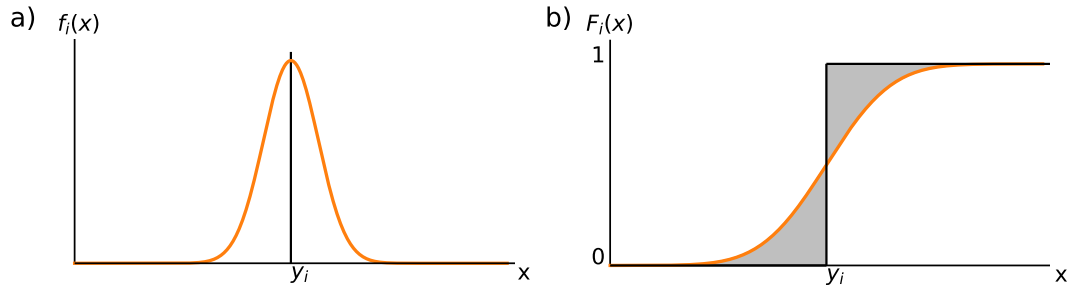



**Figure 4.2:** Illustration of the calculation of CRPS of a given a) forecast probability distribution ($f_i(x)$) (orange) and the scalar observation, $y_i$ and, b) corresponding predicted CDF (orange) and the observation CDF (black step function). The CRPS is defined as the area between these two (grey).

instantaneous continuous ranked probability score (CRPS) is determined by computing the quadratic difference between the predicted cumulative distribution function (denoted by $F$) and the observed empirical cumulative distribution function (denoted by $\mathbf{1}(x \geq y)$) (Zamo and Naveau, 2018).:

$$CRPS(F, y) = \int_{-\infty}^{\infty} [F(x) - \mathbf{1}(x \geq y)]^2 dx \tag{4.8}$$

A graphical representation of the definition of CRPS, outlined by 4.8, is displayed in Fig. 4.2. The CRPS calculation is based on a forecast probability density, exemplified in Fig. 4.2a, and determined as the grey shaded area located below the curve in Fig. 4.2b. The area is delimited by the forecast CDF (orange) and the observation CDF (black step function), with the CRPS reflecting an overall measure of forecast accuracy. The better the predicted CDF (orange) is in approximating the observed CDF (black step function), the lesser the space between them, leading to a decreased CRPS measurement. Therefore, like the Brier score, the CRPS is negatively oriented, with values ranging from 0 to $\infty$, with zero being the CRPS of a perfect forecast. Using the CRPS in forecasting has the benefit of sharing the same unit as the underlying variable being forecasted. For instance, if one predicts rainfall probability, the CRPS will have the same unit as rainfall - often mm. In situations where the forecast is deterministic, the CRPS reduces to mean absolute error (MAE), enabling direct comparison of probabilistic forecasts to deterministic ones (Gneiting and Raftery, 2007).

Like the Brier skill score, a skill score based on CRPS values can also calculated to compare the skill for a forecast relative to a reference forecast. The continuous ranked probability skill score (CRPSS) is thus defined as:

$$CRPSS = \frac{CRPS_{\text{reference}} - CRPS}{CRPS_{\text{reference}} - CRPS_{\text{perfect}}} = 1 - \frac{CRPS}{CRPS_{\text{reference}}} \tag{4.9}$$

where $CRPS_{\text{reference}}$ is the CRPS of the reference forecast and $CRPS_{\text{perfect}}$ is the CRPS of the perfect forecast, which is zero.

#### 4.5.3.5 Probability Integral Transform

The Probability Integral Transform (PIT) states that if the cumulative distribution function ($F$) of any random variable ($X$) is treated as another random variable ($Y$), i.e. $Y = F_X(X)$, then the resulting distribution of the new random variable, i.e. $F_X$, is always standard uniform (Dodge, 2003). Analysing a forecast's PIT histogram, which approaches that of a standard uniform distribution determines the forecast's degree of calibration. This principle can be applied to evaluate the calibration of probabilistic forecasts, as demonstrated in Fig. 4.3.

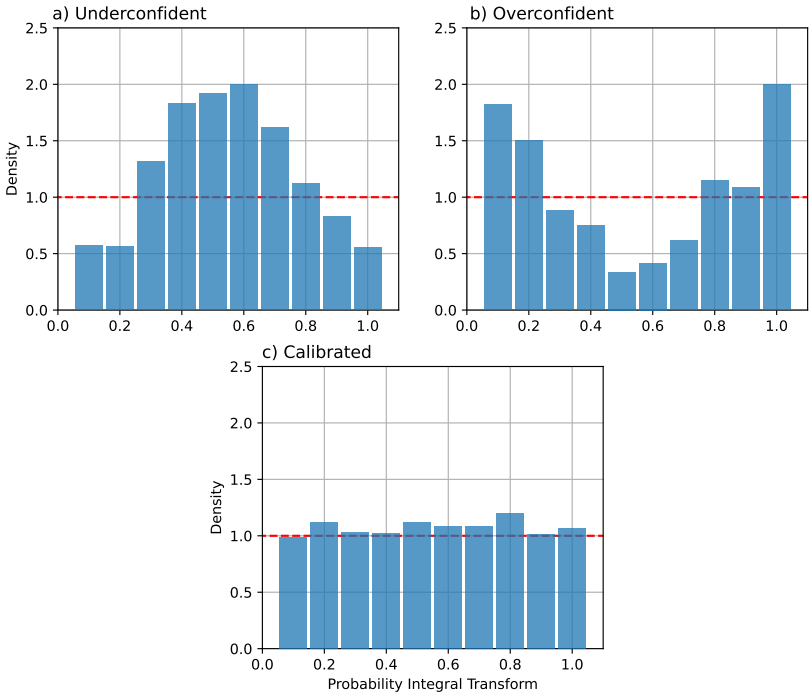



**Figure 4.3:** Illustration of Probability Integral Transform (PIT) histograms for a) underconfident, b) overconfident, and c) calibrated forecast using toy examples. The red dashed line denotes uniform distribution, indicating perfect calibration.

Deviation from uniformity in the distribution can be associated with a certain type of miscalibration (Gneiting et al., 2005). If the probability integral transform (PIT) histogram takes on a bell-like shape, the forecast can be considered underconfident (see Fig. 4.3a), while a U-shape indicates an overconfident forecast (see Fig. 4.3b). A calibrated forecast is reflected by a uniform PIT histogram, as depicted in Fig. 4.3c. For discrete-valued forecasts, such as forecasts composed of a finite number of ensembles, the PIT histogram is transformed into a 'verification rank' histogram (Anderson, 1996), where each bin represents a member of the ensemble. Nonetheless, the fundamental principle remains the same - a uniform distribution denotes calibration, facilitating comparisons between probabilistic and ensemble forecasts.

### 4.5.4 Benchmarking: EPC15-ensemble, EPC15-EasyUQ, ENS and ENS-EasyUQ forecasts (Rainfall amounts)

In order to benchmark the probabilistic rainfall accumulation forecasts, i.e. the data-driven deterministic forecasts (gamma regression and CNN) that have been run through EasyUQ (hereafter gamma regression and CNN refer to probabilistic forecasts unless explicitly stated otherwise), we compare them to the following benchmark forecasts: EPC15-ensemble, EPC15-EasyUQ, ENS and ENS-control EasyUQ forecasts.

The EPC15 ensemble forecast is a complete EPC distribution (**?**) calculated using a 15-day window on each side of the target day (31 days in total) for all days. Since we utilise observations from 2007 to 2016, each day of the EPC15-ensemble will consist of 279 members (31 multiplied by 9 training years). The EPC15-EasyUQ forecast is a probabilistic prediction that is computed by transforming the deterministic forecast, which is calculated by averaging the 15-day window in the EPC15-ensemble forecast for all days, using EasyUQ. The ENS forecast consists of the full 51-member (1 control + 50 perturbed) operational ECMWF ensemble forecast. The ENS-EasyUQ forecast is a probabilistic forecast computed by transforming the deterministic ENS control forecast through EasyUQ. We maintain the 'leave-one-out' cross-validation mode, meaning we train all benchmark forecasts (excluding the EPC15-ensemble and ENS) with data from the monsoon season (JAS) of all years except the year of the forecast and combine them by repeating the evaluation for each year from 2007-2016.

To determine if observed differences in the mean Continuous ranked probability score of the statistical and benchmark forecasts could be due to chance despite equal theoretical skill, the Diebold-Mariano test (Diebold and Mariano, 2002) is applied alongside the Benjamini-Hochberg correction (Benjamini and Hochberg, 1995), following Wilks' (2016) recommendation for the spatio-temporal setting under consideration.

# 5. Development of statistical models and new diagnostics

This chapter offers an exhaustive account of the models and diagnostic tools employed in the dissertation. They are either recently introduced to the field of meteorology or are developed for the purpose of this dissertation.

## 5.1 Statistical models

Statistical models offer an alternative to numerical weather prediction (NWP) models, which simulate the atmosphere by solving prognostic equations that describe the atmosphere given the necessary boundary conditions and the initial state. Statistical models, on the other hand, predict the outcome of target variables, such as precipitation, as a function of one or more predictor (weather) variables.

Statistical forecast models built on ordinary linear regression provide the simplest way of forecasting a target variable using a linear combination of predictor variables. This method hinges on the assumption that a predictable change in a predictor variable will subsequently result in a linearly proportional change in the target variable. However, this assumption's validity is contingent upon the target variable being equally likely to vary in either direction. Therefore, standard linear regression frequently provides inaccurate predictions of precipitation accumulation as it only differs in the positive direction, i.e., negative predictions of precipitation levels have no physical meaning.

### 5.1.1 Generalised linear models

Generalised linear models (GLMs) (Nelder and Wedderburn, 1972) overcome the limitation of standard linear regression models by enabling the target variables to be derived from arbitrary distributions. GLMs achieve this by using a 'link' function $h$, of the target variable $\hat{y}$ (in the case of ordinary linear regression models, the link function is the Identity function) to vary linearly with the predictor variables $X$ (as linear combinations of unknown parameters $\omega$ and predictor variables $X$), as opposed to varying $\hat{y}$ itself. The relationship between the $\hat{y}$ and $h$ is as follows:

$$\hat{y}(\omega, X) = h(X\omega), \tag{5.1}$$

Following this approach, the traditional minimization of a squared loss function in ordinary linear regression is replaced with the following:

$$\min_{\omega} \frac{1}{2n_{\text{samples}}} \sum_{i} d(y_i, \hat{y}) + \frac{\alpha}{2}||\omega||_2^2, \tag{5.2}$$

where $\alpha$ is a constant that multiplies the $L2$ penalty term, determining the strength of the regularization, while $d$ is defined as the unit deviance and $y_i$ are the observations.

#### 5.1.1.1 Logistic regression model

The logistic regression model is a particular case of GLMs with a logistic distribution (also referred to as the inverse *logit* function), typically used for forecasting the probability of events with binary outcomes, which, in the context of this thesis, is rainfall occurrence (McCulloch, 2000). Following Vogel et al. (2021), we now construct a purely data-driven, statistical logistic regression (hereafter, Logistic) model. To this end, let $o_1$, $o_2$, and $o_3$ denote the observed precipitation accumulations at the grid points of the highest Coefficient of Predictive Ability (CPA, see section 5.2.1) at lags of one, two, and three days, respectively. Then the logistic model for the probability $p$ of 24-hour precipitation occurrence at the target gridpoint depends on $o_1$, $o_2$, $o_3$ and the day of the year $D$, positing that

$$\log \frac{p}{1-p} = a_0 + a_1 f(o_1) + a_2 f(o_2) + a_3 f(o_3) + b_1 \sin \frac{2\pi D}{365} + b_2 \cos \frac{2\pi D}{365}, \tag{5.3}$$

where the final two terms depend periodically on the day of the year. The function $f(x) = \log(x + 0.001)$ is used to transform a non-negative rainfall amount into a real value. We estimate the statistical parameters $a_0$, $a_1$, $a_2$, $a_3$, $b_1$, and $b_2$ in (5.3) using the LogisticRegression subroutine from the scikit-learn package (Pedregosa et al., 2011a). The parameters are estimated for each target grid point individually, and so they vary spatially. Following Vogel et al. (2021), we operate in cross-validation mode, i.e., the logistic model is trained on data from all years available except the year for which the prediction is issued. This is repeated until all the years from 2007 to 2019 are used. 'Leave-one-out' cross-validation makes the best use of the available data and limits artefacts due to a low number of samples, thereby producing more robust results. As a consequence of this approach, we developed a separate model for every year from 2007 to 2019.

#### 5.1.1.2 Gamma regression model

As noted earlier, rainfall accumulation follows a non-negative distribution and is often associated with a long right tail, as extreme values are rare and anomalous. To account for these factors, we utilise a gamma-distributed GLM (hereafter gamma regression)[1]

[1]To implement gamma regression, we use the `GammaRegressor` function from `scikit-learn` (Pedregosa et al., 2011b)

where $d$ in Eq. 5.2 is defined as

$$d(y, \hat{y}) = 2\left(log\frac{y}{\hat{y}} + \frac{y}{\hat{y}} - 1\right), \tag{5.4}$$

as given by Jørgensen (1987). However, since the gamma distribution can only be fitted to positive values, we set all non-positive values in $y$ to a small positive quantity $\delta$, similar to the approach explained in Scheuerer and Hamill (2015), as follows:

$$y' = \begin{cases} y & y > 0 \\ \delta & y \leq 0, \end{cases} \tag{5.5}$$

The steps involved in the gamma regression model are as follows: 1) Fit the model on the training set with $\alpha = 0.0$ for a maximum of 1 million iterations or until the minimum loss is achieved. We use the L-BFGS-B algorithm (Zhu et al., 1997) to achieve convergence; 2) Generate predictions using the test data on the trained model. Since we operate in 'leave-one-out' cross-validation mode, these steps are repeated for every train-test split, and the results are concatenated to obtain the full deterministic forecast for the 2007-2016 period.

### 5.1.2 Convolutional neural-network model

While GLMs, such as the gamma regression model, are capable of learning 'basic' non-linear relationships between predictor and target variables, their ability remains doubtful when handling a significant amount of predictor variables. Artificial Neural Networks (ANNs), also known as Neural Networks (NNs), are machine learning algorithms that can model more complex non-linear relationships between predictor and target variables in an effective manner (Nielsen, 2015). Convolutional Neural Networks (CNNs) are a particular category of ANNs designed and frequently used to analyse and process structured, gridded data. It is a type of 'regularized' neural network that operates using feed-forward data flow. The implementation of 'convolutions' in this network makes it less prone to overfitting than other ANNs. Convolutions connect two sets of data and are used to filter information (known as a 'kernel') from the input data to create a feature map. The kernel is then slid over the data until the map is full. This method offers certain benefits, including greater computational efficiency and improved accuracy compared to ANNs.

Although convolutional neural networks (CNNs) are extensively employed in image processing and classification, recent studies have shown that they can be effectively applied to the prediction of sequential data (e.g., Bai et al., 2018; Tsantekidis et al., 2017) in which causality is important (e.g., time series data). Likewise, Liu et al. (2016) have utilised CNNs in detecting extreme weather phenomena such as tropical cyclones.

The process of prediction using the CNN model involves the following steps: 1) To eliminate any imbalances in the dataset, e.g., due to different orders of magnitudes of

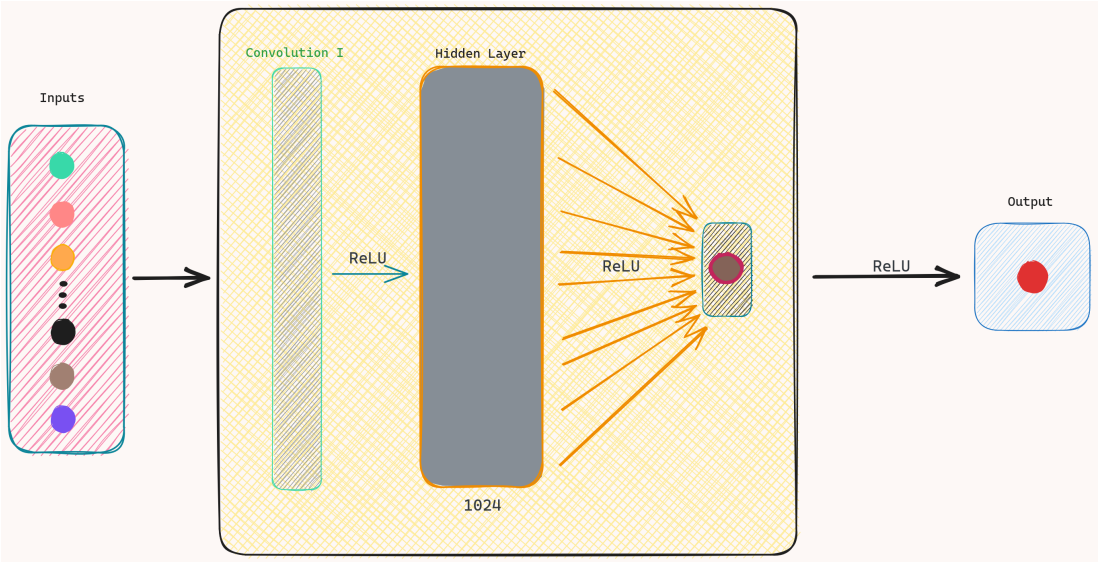



**Figure 5.1:** The architecture of the Convolutional Neural Network (CNN) model used in the study is illustrated. The data flow is from left to right. Inputs first pass through a causally padded convolutional layer with 32 filters of kernel size 3. Non-linearity in this layer is induced by a Rectified Linear Unit (ReLU) activation function, which acts like a filter setting all negative values to zero. The output of the convolutional layer is then passed through a hidden layer of 1024 ReLU-activated neurons, the output of which is then condensed to a single value by a single neuron output layer, also activated with a ReLU activation function. Finally, in order to obtain a non-negative precipitation value, the output layer is also activated with ReLU.

predictor and/or target variables, we apply scaling to the predictor and target variables. We achieve this by dividing each variable by its maximum absolute value, thereby ensuring that the maximum absolute value in the dataset is 1.0 without centering or shifting the data. 2) We further partition the training data into a training set and a validation set, using the last season in the training set for validation. 3) The CNN[2] model (refer to Figure 5.1 for CNN model architecture details) undergoes 100 epochs of training employing a learning rate of 0.0001. The training employs early stopping if the evaluation metric's validation score (refer to section 4.5 for details) does not indicate an improvement surpassing 5 epochs. 4) The 'Adam' optimizer is utilized during model fitting due to its computational efficiency and lower memory consumption in comparison to other stochastic gradient descent algorithms (Kingma and Ba, 2014). 5) Subsequently, predictions are generated using the test data and the results are rescaled to their original range. These steps are repeated for every train-test split, and the outcomes are concatenated to obtain the full deterministic forecast for the entire 2007-2016 period, as we are operating in cross-validation mode.

## 5.2 New diagnostics

For the purpose of selecting predictors for training the statistical models and, more significantly, supporting the importance of chosen predictor variables using meteorologi-

[2]To implement CNN, we use the `conv1D` function and associated functions within the `keras` Python package (Chollet et al., 2015)

cal reasoning, numerous innovative diagnostics have been employed. These diagnostics have either been recently introduced in the field of meteorology or custom-developed specifically for this dissertation.

### 5.2.1 Selection of predictor variables via CPA

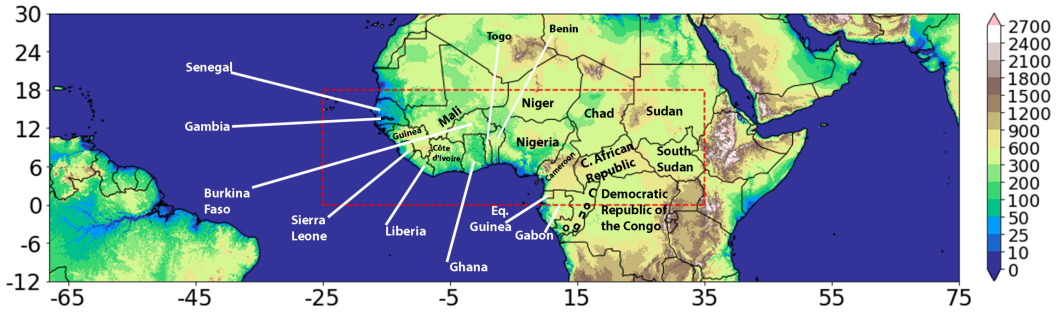



**Figure 5.2:** Overview of the study domain. We restrict the development of the purely data-driven statistical model and its evaluation to the nested core area that comprises 25° W to 35° E in longitude and 0° to 18° N in latitude. The shading indicates altitude in meters. Reprinted from Rasheeda Satheesh et al. (2023). ©2023 The Authors, John Wiley & Sons Ltd.

To identify predictor variables for our purely data-driven statistical model, we use 0600 UTC to 0600 UTC accumulated GPM IMERG precipitation within the nested core area (Fig. 5.2) as target data, with the same field over the full area at temporal lags of one, two, and three days serving to provide candidate variables for the statistical model. For each target grid point, we compute the Coefficient of Predictive Ability (CPA) measure between the local precipitation accumulation and all of the candidate variables.

The recently developed CPA measure is tailored to the selection of predictor variables for forecasting mixed discrete-continuous variables such as accumulated precipitation (Gneiting and Walz, 2022). CPA is asymmetric and, thus, takes the distinct roles of the target variable – here, 24-hour precipitation accumulation at a target grid point ($B$) – and the potential predictor variables into account ($A$), i.e., in general,

$$\mathrm{CPA}(A,B) \neq \mathrm{CPA}(B,A). \tag{5.6}$$

For data without ties, CPA becomes symmetric and can be expressed in terms of Spearman's rank correlation coefficient $\rho_S$, which was used by Vogel et al. (2021), namely, as

$$\mathrm{CPA} = \frac{\rho_S + 1}{2}. \tag{5.7}$$

While precipitation accumulations tend to show ties, this relationship only holds to a good degree of approximation over most of our analysis region. If the outcome is binary, CPA reduces to the classical Area Under the Receiver Operating Characteristic (ROC) curve (AUC) measure. Like the AUC measure, CPA is a non-negative quantity that is bound between 0 and 1, with a value of 0.50 indicating the absence of a predictive skill,

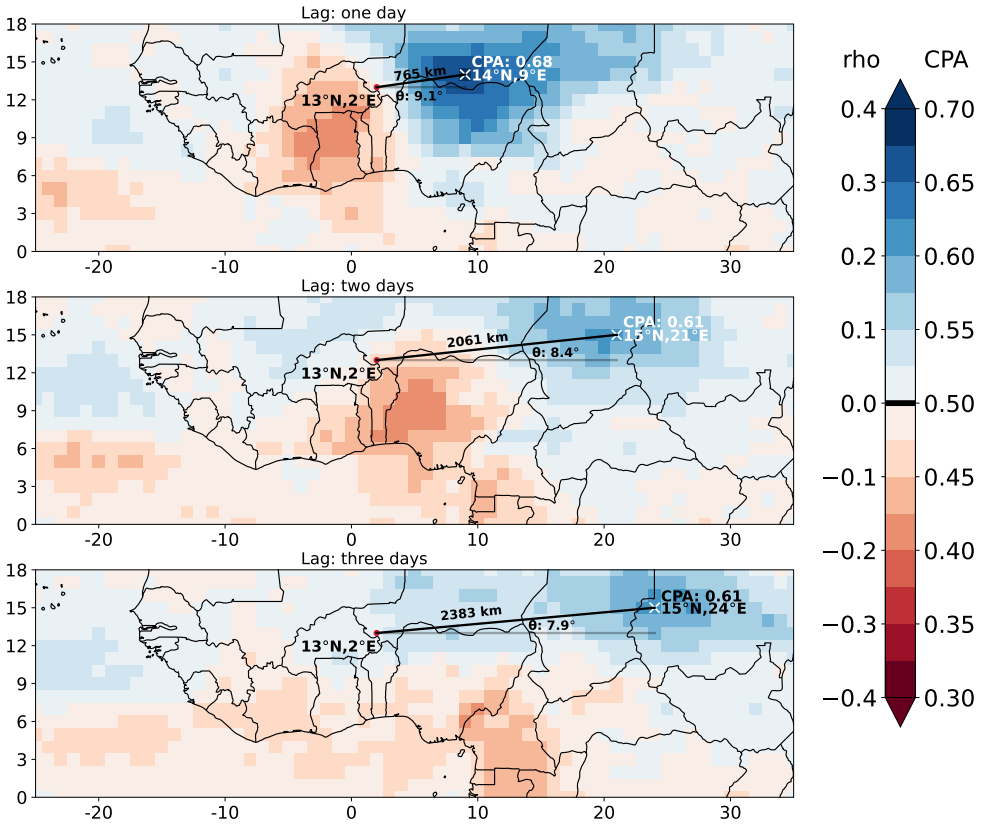



**Figure 5.3:** CPA with respect to the grid point (marked by a red dot) closest to Niamey (13° N, 2° E) at temporal lags of (a) one, (b) two, and (c) three days, based on data from July–September 2001–2019, except for 2016. The white cross (×) marks the grid point of maximum CPA at the given lag, with distance and direction ($\theta$) from Niamey. The shading can also be interpreted in terms of the approximate value of Spearman's rank correlation $\rho_S$ according to Equation 5.7. Reprinted from Rasheeda Satheesh et al. (2023). ©2023 The Authors, John Wiley & Sons Ltd.

and a value of 1 representing a deterministic, strictly increasing relationship. Values of CPA below 0.50 correspond to decreasing relationships and negative correlations.

As we operate in cross validation mode – with data from a single year left out, and the rest of the data used to fit the statistical model – data from the season in consideration for the forecast are not included. For example, when developing the forecast model for the year 2016, data from 2016 are left out, but data from all other years between 2001 and 2019 are used, and this is repeated for all years from 2001 to 2019. As an illustrative example for the CPA analysis, Fig. 5.3 considers the target grid point closest to Niamey (Niger), with data from 2016 left out. At a lag of one day, the maximum arises 765 km to the east of Niamey, reaching a CPA value of 0.68. At a lag of two days, CPA generally decreases in magnitude, and grid points with high CPA are shifted away from Niamey. At a lag of three days, the maximum CPA of 0.61 is reached at the border of Chad and Sudan nearly 2400 km east of Niamey in roughly the same direction as the other lags. The results are indicative of westward propagation with an approximate phase speed of 8.5 $ms^{-1}$, which is typical for rainfall affected by AEWs.

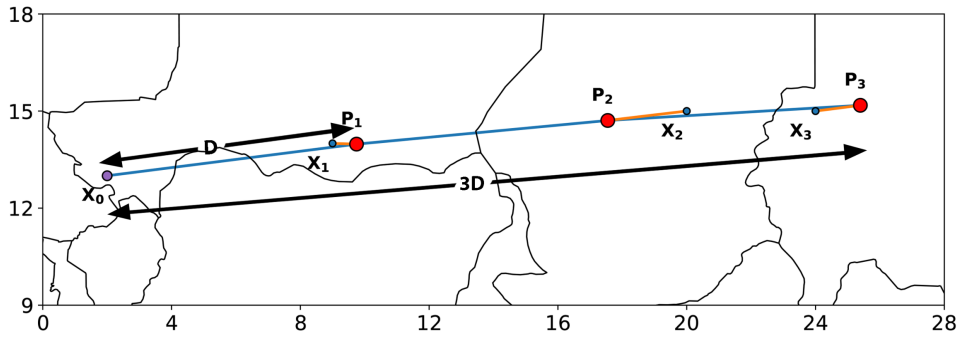



**Figure 5.4:** Computation of the coherent-linear-propagation factor (*coh*) at the grid point $X_0$ closest to Niamey (13° N, 2° E). The grid points $X_1$, $X_2$, and $X_3$ (blue dots) are the locations of maximum CPA with respect to $X_0$ at lags of one, two, and three days, respectively, based on data from July–September 2001–2019 except for 2016. The points $P_1$ and $P_2$ lie on the great circle segment from $X_0$ to $P_3$ and mark the corresponding ideal locations of maximum CPA in case of a perfectly linear propagation with constant phase speed. The grid point $P_3$ depicted here is chosen such that the error term in Equation 5.8 (orange lines) becomes minimal. Reprinted from Rasheeda Satheesh et al. (2023). ©2023 The Authors, John Wiley & Sons Ltd.

These patterns are exploitable for statistical forecasting and resemble those in Fig. 1 of Vogel et al. (2021), suggesting good consistency between TRMM and IMERG data.

### 5.2.2 Coherent-linear-propagation factor

To quantify the coherence of the patterns identified by the above analysis, we define a coherent-linear-propagation factor (*coh*) at any given target grid point. Figure 5.4 illustrates the concept on the same grid point as in Fig. 5.3. Specifically, let $X_0$ be the target grid point under consideration, and let $X_1$, $X_2$, and $X_3$ (blue dots in Fig. 5.4) be the grid points of maximum CPA at lags of one, two, and three days, respectively. We aim to quantify the extent to which the positions of these points agree with propagation at constant speed and constant direction. To this end, we introduce auxiliary, equispaced points $P_1$, $P_2$, and $P_3$ on a great circle segment that starts from $X_0$ with distance $D$ between successive points (Fig. 5.4). This way, the great circle distance (gcd) between $X_0$ and $P_3$ equals $3D$. The deviation from a linear propagation can now be quantified by the error term (orange lines in Fig. 5.4)

$$E = \Big( \mathrm{gcd}(X_1, P_1)^2 + \mathrm{gcd}(X_2, P_2)^2 + \mathrm{gcd}(X_3, P_3)^2 \Big)^{1/2}. \tag{5.8}$$

Analogous to a regression problem, we can now find the optimal choice of the grid point $P_3$ at which $E$ achieves a minimum. We denote this minimum by $E_0$, and we let $D_0$ be one-third of the great circle distance between $X_0$ and the optimal $P_3$.

The magnitude of $E_0$ will tend to be larger for grid points showing fast propagation, i.e., for large $D_0$, such that these quantities should be considered in relation to each

other. We therefore define[3] the *coh* at the target grid point $X_0$ as

$$coh = \frac{3D_0}{3D_0 + E_0}. \tag{5.9}$$

In Fig. 5.4, the grid point at $P_3$ is such that the error term in (5.8) is minimized. The *coh* measure then equals the ratio of the length of the blue great circle segment to the sum of the length of the blue segment and the error term; specifically, $D_0 = 844.4$ km, $E_0 = 315.8$ km, and coh = 0.889. Generally, under near-perfect linear propagation, $E_0$ will be small and thus, *coh* will be close to its ideal value of 1. If the points of maximum CPA scatter at random, $E_0$ will be large and *coh* will be small (but remain non-negative). In a nutshell, high values of the *coh* measure characterize regions dominated by stable propagation of long-lived atmospheric features that modulate conditions for rainfall, implying comparably high predictability that could potentially be exploited with statistical models. Conversely, low *coh* characterizes regions that do not experience stable, coherent linear propagation. For these regions, rainfall might be largely stochastic or might be subject to distinct forcings. It is important to note that these characterizations are in terms of linear propagation; more complex methods might be able to detect acceleration and deceleration as well as nonlinear tracks. Also, the definition is only sensitive to geographic distances and not the values of CPA itself, which of course, do matter for the skill of statistical models built on these relationships.

### 5.2.3 Computation of local phases and amplitudes of tropical waves

To compute local predictors based on TWs, which are planetary-scale features, we adopt a modified version of the phase-amplitude computation outlined by Yasunaga and Mapes (2012) and van der Linden et al. (2016). To summarize: a) For any given wave at any given location (or averaged over an area), compute the filtered wave signal and its time derivative, and normalize them by dividing by the standard deviation; b) Compute a phase-amplitude diagram as shown in Fig. 5.5 where the $x$ and $y$ axes are the filtered wave signal and its time derivative respectively; c) The radial distance of a scatter-point in this setup is defined as the amplitude at that point in time and the angle subtended by that point with the positive $x$ axis is defined as $\theta$. Furthermore, the phase-amplitude diagram is divided into eight sectors (or phases) named from P1 through P8 (from now on, referred to as simply 1–8). Phases 4–6 are defined as the local 'wet' phases (positive rainfall anomaly), and phases 8–2 are the local 'dry' phases (negative rainfall anomaly). Phases 3 and 8 are defined as the 'neutral' or 'transition' phases where the wave is in its neutral state or the wave is transitioning between wet and dry phases.

The categorization of timesteps into the eight phases will be used in Chapter 7 to investigate the relationship between the passage of wave modes and the skill of rainfall forecasts. For Chapter 8 that focuses on forecasting rainfall accumulation using TWs

[3] If $X_0 = X_1 = X_2 = X_3$, we note a lack of propagating features at the target grid point at $X_0$, and thus we set coh = 0. In all other cases we proceed as stated.

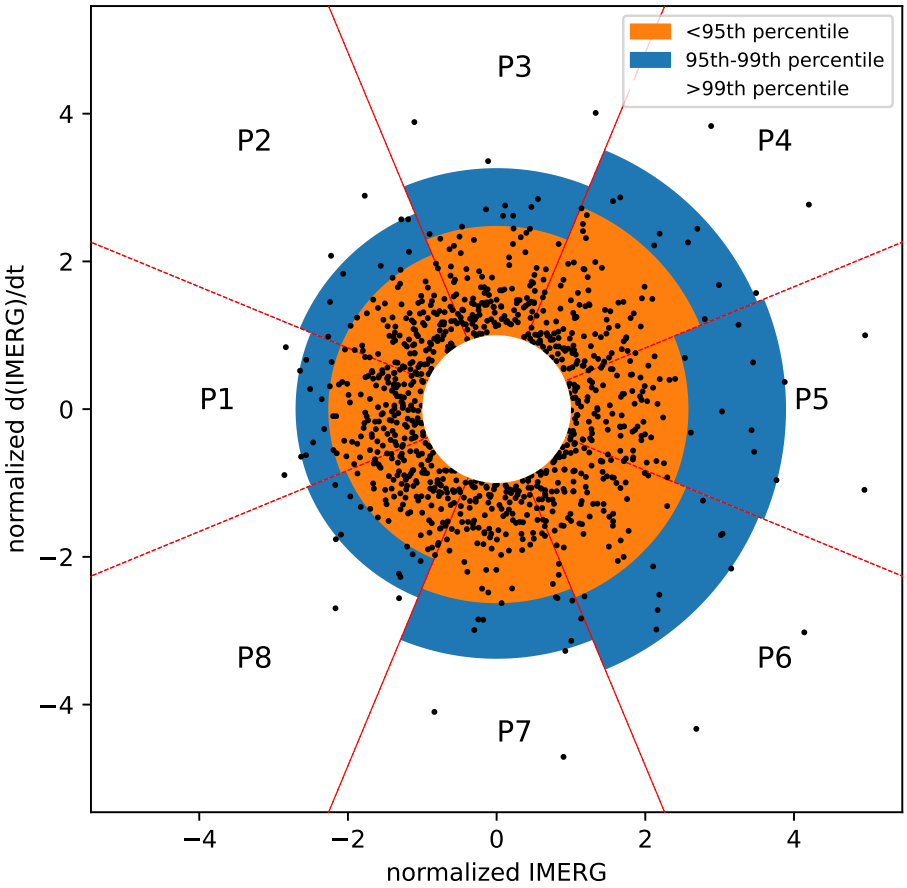



**Figure 5.5:** The ER wave's local phase-amplitude diagram at a gridpoint near Niamey (13° N, 2° E) obtained using normalized IMERG rainfall is shown. The orange and blue shadings represent amplitudes that fall below the 95$^{th}$ percentile and between the 95$^{th}$ and 99$^{th}$ percentiles, respectively. Points that lie outside these regions have amplitudes above the 99$^{th}$ percentile. The white circle at the centre masks amplitudes below 1 standard deviation. The phase-amplitude diagram is divided into eight sectors, named from phases P1 through P8, based on the 'wetness' of the wave activity (refer to section 5.2.3 for details).

as predictors, the phase information (whether wet or dry) and the amplitude (local strength of the wave) are combined to produce the 'phase-adjusted wave amplitude' (hereafter, PWA) as follows for any given wave at timestep $t$:

$$\mathrm{PWA(wave)_t} = \mathrm{amplitude(wave)_t} \times \cos\theta\mathrm{(wave)_t} \tag{5.10}$$

To ensure continuous values in the predictor variables and to use all the available data, the computation of PWA takes into account all timesteps (including the amplitudes that fall below 1.0 standard deviation).

### 5.2.4 Tropical wave-based predictor selection via Gradient-Boosting

To identify predictor variables based on tropical waves, we utilise the local phases and amplitudes of TWs (PWA) at each grid point, as described in the previous section. To

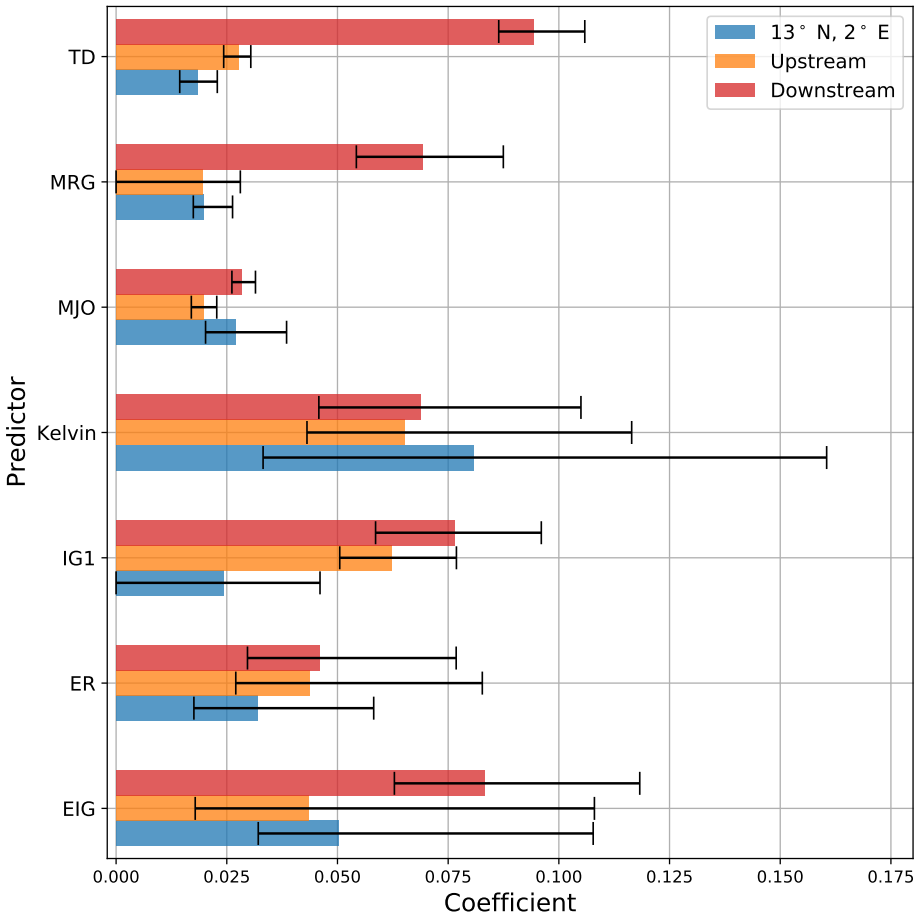



**Figure 5.6:** Mean values of feature importance of predictor variables at the grid point nearest to Niamey (13° N, 2° E) for every training set from 2007-2016. The blue bars represent average feature importance at (13° N, 2° E), while red bars represent average feature importance one grid point downstream of TWs propagation (i.e., 13° N, 3° E for westerly waves and 13° N, 1° E for easterly waves) and the orange bars represent the average feature importance one grid point upstream of TWs (i.e., 13° N, 1° E for westerly waves and 13° N, 3° E for easterly waves). The error bars represent the range of coefficients in each iteration of train-test splits from 2007-2016.

forecast rainfall accumulation at a specific grid point, we calculate predictor variables at three grid points: the target grid point (i.e., where the prediction is issued), one grid point upstream, and one grid point downstream in the direction of propagation of the TW. Since we are interested only in the zonal propagation of the TW, the upstream and downstream grid points lie along the same latitude as the target. This process is repeated for all seven TWs. Therefore, for training a forecast model at the target grid point, we have a total of 21 potential predictor variables.

A potential drawback of utilizing predictor selection via CPA (outlined in Section 5.2.1) is its complete reliance on human expertise. While it suffices in its current context, with only three predictors selected for each target grid point, implementing this approach for 21 potential predictors would be exceedingly laborious and possibly error–prone. Hence, we implement a semi–automated feature selection procedure to select the most effective

predictor variables at each grid point via gradient boosting[4] (Friedman, 2001; Chen and Guestrin, 2016). The feature selection procedure is structured as follows: 1) The training dataset is partitioned into two subsets, whereby the final year of the training dataset is assigned for validation; 2) The gradient boosting model is fitted, and the feature importance[5] of the predictor variables that minimise the loss in the validation set is calculated. We assume that the daily rainfall accumulation follows a gamma distribution for the same reasons described in Section 5.1.1.2. Therefore, we opted for the 'reg:gamma' option to calculate the output loss, which reflects the mean of the gamma distribution; 3) We then proceed to select those predictor variables that exceed a threshold of $0.75 \times median$ of the feature importance values as it produced the best results in our model setup. Note that since we operate in 'leave-one-out' cross-validation mode, these procedures are repeated for every train-test split from 2007-2016. To illustrate the outcomes of this approach, the mean values of the feature importances and their ranges from 2007-2016 at the grid point near Niamey (13° N, 2° E) are shown in Fig. 5.6.

Interestingly, the wave activity of tropical depression (TD) at the downstream grid point (13° N, 3° E) was found to be a more reliable predictor compared to that at the target grid point (13° N, 2° E). In order to clarify this, we hypothesize the following mechanism: Suppose that precipitation, which was modulated by a tropical wave like the westward propagating TD, occurred simultaneously at 00 UTC in both the downstream and target grid points (13° N, 2° E). As it is unlikely for the TDs to convey this data to the target location as the wave propagates further away from the target grid point with time, the quantity of rainfall impacted by TD at the downstream location plays a crucial role in forecasting rainfall at the target grid point, albeit in a negative way. In other words, the information about rainfall in the downstream gridpoint ensures that rainfall occurrence at the target grid point is unlikely in the next 24 hours. The same mechanism applies to the rainfall on the target grid point as well, but it bears little importance as compared to the downstream grid point. This issue is elaborated upon in detail by Lafore et al. (2017) and is identified as the primary cause of low positive to even negative autocorrelation of precipitation anomalies at the 1-day lag over tropical Africa as shown in Roehrig et al. (2013). The same mechanism can be used in cases when the upstream grid point is preferred, i.e., the precipitation data at the upstream grid point holds greater significance than the target location as it provides information about incoming precipitation systems.

The narrower error bars on the downstream (red bars in Figure 5.6) and upstream predictors (orange bars in Figure 5.6) for TDs provide additional evidence for the potential robustness of this mechanism. Similar patterns are also noticeable in the Kelvin and IG1 waves, indicating appropriate strengths. This suggested mechanism is applicable

[4]Gradient boosting is implemented using the `eXtreme Gradient Boosting (XGBoost version 1.7.6)` Python package (Friedman, 2001; Chen and Guestrin, 2016).

[5]Defined as 'gain' in `XGBoost`

to all types of wave modes without regard to their direction. Another noteworthy finding is that for all westward-moving waves, predictors are ranked in the order of $downstream > upstream > target$, except for the MRG wave. However, it remains unclear why the observed pattern for eastward-propagating waves is $downstream > target > upstream$ instead, with the exception of the Kelvin wave.

# 6. Statistical forecasting of 24-hour rainfall occurrence using Logistic regression

The study presented here builds on the foundations laid by Vogel et al. (2021) and develops them in several ways: We strengthen the purely data-driven statistical model for 24-hour precipitation occurrence by (a) using the recently developed Coefficient of Predictive Ability (CPA; see section 5.2.1) for predictor selection and, (b) expanding the logistic regression model to consider temporal lags of up to three days.[1] The CPA bridges Spearman's rank correlation coefficient, as used by Vogel et al. (2021), and the Area Under the Receiver Operating Characteristic (ROC) curve (AUC) in a way that makes CPA a non-negative asymmetric quantity bound between 0 and 1 (Gneiting and Walz, 2022). Moreover, (c) we introduce the coherent-linear-propagation factor as a metric that quantifies the coherence of a propagating rainfall signal identified over lags of one, two, and three days, (d) we provide – at least qualitatively – physical interpretations of the most significant patterns using auxiliary meteorological fields and, (e) we replace the standard analysis of Brier scores and reliability diagrams with more robust methods, namely, the recently developed CORP (see section 4.5.1.2 ) approach (Dimitriadis et al., 2021). Finally, as minor updates, (f) we now use the newer GPM IMERG product instead of the discontinued Tropical Rainfall Measuring Mission (TRMM) data, and (g) the study domain has been extended significantly beyond West Africa, now including various environments (e.g., ocean, flatlands, highlands) in order to facilitate the identification of coherent propagating features. The ultimate goal of our work is to better understand the potential sources of predictability related to the time-space behaviour of rainfall in tropical Africa and to test the skill of the purely data-driven statistical model against the latest benchmarks with respect to predicting precipitation occurrence in the next 24 hours.

The chapter is structured as follows. In section 6.1, spatio-temporal patterns, the direction and speed of propagating rainfall, and the degree of coherence are discussed for the summer monsoon season. Following this, in section 6.2, we compare the predictive

[1] Three days appears to be a good compromise between including all relevant information and keeping the domain size and computations at a reasonable level.

performance of the purely data-driven logistic regression model to a climatological forecast (termed EPC15), the European Centre for Medium-Range Weather Forecasts (ECMWF) raw ensemble (ENS), and a statistically post-processed ensemble forecast (ENS-IDR). We repeat the same analyses for the pre-monsoon (May-June; MJ) and the two transition seasons (March-April; MA and October-November; ON) and discuss the highlights in section 6.3. We do not consider the winter season as most of tropical Africa remains dry. The chapter closes with Section 6.4, where we summarize and discuss the major findings of this study.

## 6.1 Maximum CPA and coherent-linear-propagation factor

Using the methodology described in Section 5.2.1, Fig. 6.1 provides a detailed analysis of the points of maximum CPA relative to the target grid points in the evaluation domain.

The maximum CPA (Fig. 6.1a–c) peaks at a temporal lag of one day and for target grid points over the northern tropical Atlantic Ocean, where it reaches values of about 0.75. The African rainbelt (Nicholson, 2008, 2009) is broadly co-located with the lowest values of maximum CPA, albeit with some east–west gradient (see dark red isohyets in Fig. 6.1a–c). The latitudinal position of its axis varies from 6° N in central Africa to 10° N over West Africa, and monthly rainfall amounts range from well above 400 mm per month over the warm waters of the Atlantic to below 200 mm per month over South Sudan. Also, note the slight orographic enhancement on the western sides of the Guinea Highlands and the Cameroon Line Mountains. High values of maximum CPA are generally found in areas of large climatological rainfall gradients straddling the rainbelt. On the southern side, the zone of high CPA at a lag of one day stretches from the eastern Atlantic at about 5°N past the Guinea Coast towards southern Nigeria, Cameroon, and Gabon. Interestingly, the distortion of the rainfall gradient over the relatively drier Ivory Coast and Ghana is also reflected in a northward excursion of higher CPA values. To the north of the rainbelt, over the Sahel, CPA also is enhanced. A possible interpretation of these patterns is that the rainbelt is characterized by favourable conditions for convection, which creates a high degree of stochasticity, whereas the fringes are more sensitive to coherent synoptic-scale features that trigger or organise convection, such as slow, westward propagating cyclonic vortices that persist for more than a day, or nighttime systems that last across the 0600 UTC dividing time (Knippertz et al., 2017; Kniffka et al., 2020). Comparing results for a lag of one day with those for lags of two and three days, we find good agreement in the geographic patterns of maximum CPA despite a considerable reduction in the values (as low as 0.50).

Considering the distance (Fig. 6.1d–f) and direction (Fig. 6.1g–i) of the point with maximum CPA relative to the target grid point, we leave grid points with distances below 400 km blank, to signal very slow propagation. The largest distances are generally found for points in the continent's interior along the African rainbelt. Two types of behaviour can be distinguished. In a broad region stretching from the West Coast to

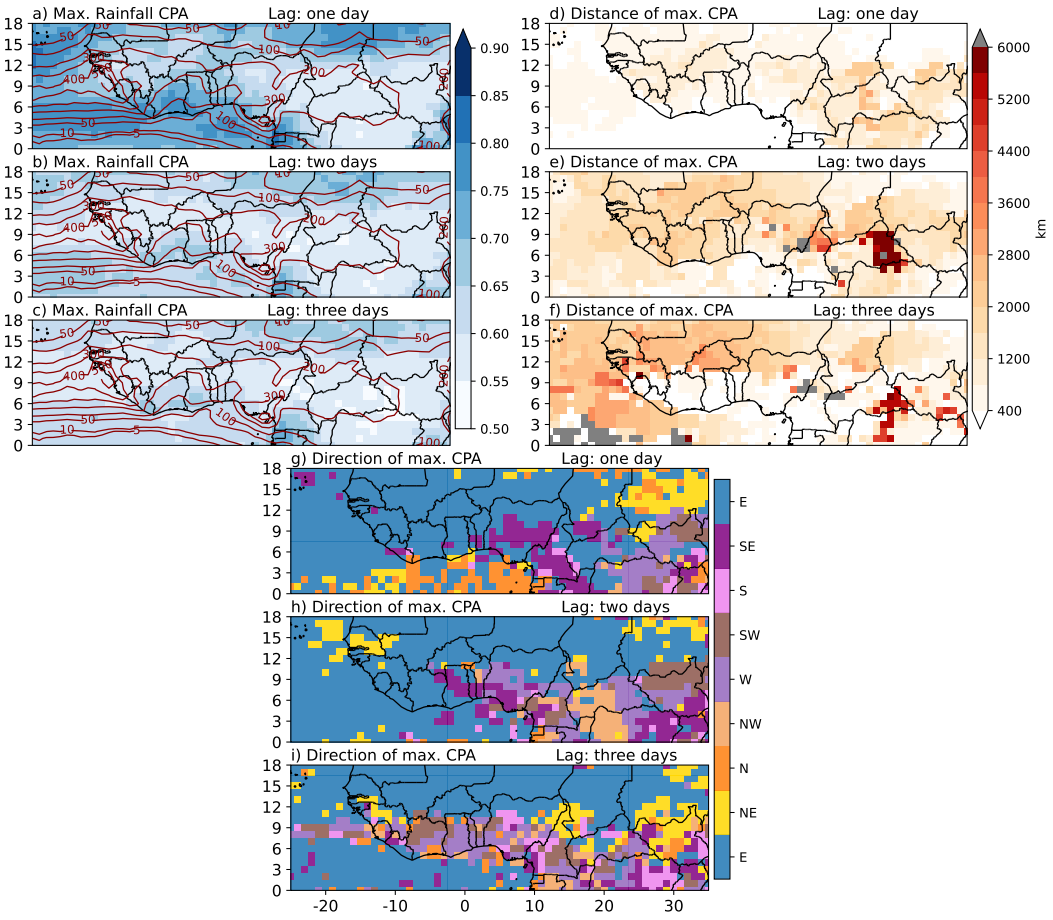



**Figure 6.1:** Maximum CPA (a–c) along with distance (d–f) and direction (g–i) between the target grid point and the grid point with maximum CPA, at temporal lags of one (a,d,g), two (b,e,h), and three (c,f,i) days, respectively, for JAS 2001–2019 except 2016. The dark red contours in (a–c) show monthly average rainfall in the period considered in mm. Reprinted from Rasheeda Satheesh et al. (2023). ©2023 The Authors, John Wiley & Sons Ltd.

20°E propagation is mostly westward (occasionally northwestward) with distances at a lag of one day, typically at 600–1000 km, which corresponds to about 6–10° longitude per day or a phase speed of about 6.9–11.5 $ms^{-1}$. This type of behaviour is also evident in the example from Fig. 5.3. It extends out to the downstream Atlantic, with propagation distances in the lower end of the range. The second type occurs over Central Africa and is characterized by mostly eastward propagation over larger distances.

We proceed to an interpretation of these results via the coh and auxiliary meteorological fields, such as vertical shear, CIN, and CAPE, in Fig. 6.2. The highest values of coh are found in the western Sahel north of 10°N, where coh reaches an impressive 0.90 over Burkina Faso, western Niger, and northwestern Nigeria (Fig. 6.2a). In view of the relatively low maximum CPA of around 0.65 at a lag of one day, this insinuates relatively fast, westwards propagating rainfall features with phase speeds in the range 7.5–12 $ms^{-1}$, in concert with smaller-scale variation, possibly weakening the spatio-temporal relationships. This aspect is also evident in Fig. 5.3, where high values of CPA are

spread out over a relatively large area. These results are consistent with large easterly vertical wind shear between 600 and 925 hPa (arrows) that is key to the formation of MCSs (Maranan et al., 2018). Given mean propagation speeds of 9 $ms^{-1}$ for AEWs and 15 $ms^{-1}$ for MCSs (Fink and Reiner, 2003), the high coherence might stem from a dominating influence of AEW propagation, which is known to strongly modulate rainfall in this area (Schlueter et al., 2019b). At the same time, shorter-lived MCSs form within the AEW envelope but propagate much faster due to the high shear, leading to relatively low CPA, as noted.

Another relevant factor in explaining the high coherence over the western Sahel is CIN (Fig. 6.2b), which shows high values throughout, with a marked south-to-north increase with values ranging from 80 $Jkg^{-1}$ over southern Mali to over 400 $Jkg^{-1}$ over northern Mali and Niger. High CIN implies that relatively strong triggers are needed to generate rainfall, e.g., associated with AEWs and MCSs. Combined with high CAPE (Fig. 6.2c), long-lived rainfall events tend to be relatively infrequent but, if occurring, rather intense, resulting in relatively low CPA and moderate overall rainfall in the western Sahel.

In the eastern Sahel, values of coh drop below 0.50, despite CPA values being similar to the west Sahel region and even higher in northern parts (Fig. 6.1a–c ). This is likely due to weaker AEWs, which grow in amplitude while propagating westward across the Sahel along the African Easterly Jet (AEJ), which is climatologically weaker in the east (Fink and Reiner, 2003; Lafore et al., 2017). Shear also is weaker than in the western Sahel and less homogeneous in direction and magnitude. Further reasons for lower coh may lie in higher CIN and smaller CAPE, which make it harder to trigger convection, as also reflected in lower rainfall (dark red isohyets in Fig. 6.1a–c).

A sharp drop in coh is observed as we move southward from the Sahel towards the centre of the rainbelt, where values of coh are generally very low. An exception is Central Africa (eastern Central African Republic and northern Democratic Republic of the Congo), where coh is locally much higher, though in combination with CPA below 0.575, as indicated by stippling in Fig. 6.2a–c, which suggests possible artefacts. Indeed, there is no dominant propagation direction in this area (Fig. 6.1g–i). Generally, westward propagation dominates in the rainbelt (Fig. 6.1g–i) combined with low CIN and high CAPE (Fig. 6.2b–c) such that convection can be triggered with ease almost anywhere, inducing high area-averaged rainfall and a high degree of stochasticity.

To the southwest of the rainbelt, there is a mostly oceanic region (tropical eastern Atlantic) with moderate to large values of coh, where values of CPA at one day lag peak around 0.75 near the coast of Liberia and Sierra Leone (Fig. 6.1a). Since convective maxima over the ocean tend to occur in the early morning (Albright et al., 1985; Hendon and Woodberry, 1993), it is conceivable that in some cases, the spatio-temporal relationship is caused by a convective storm surviving until the next morning and therefore affecting consecutive 24-hour periods, as corroborated by distances mostly below 400 km (Fig. 6.1d) and variable directions from northerly to easterly. Interestingly,

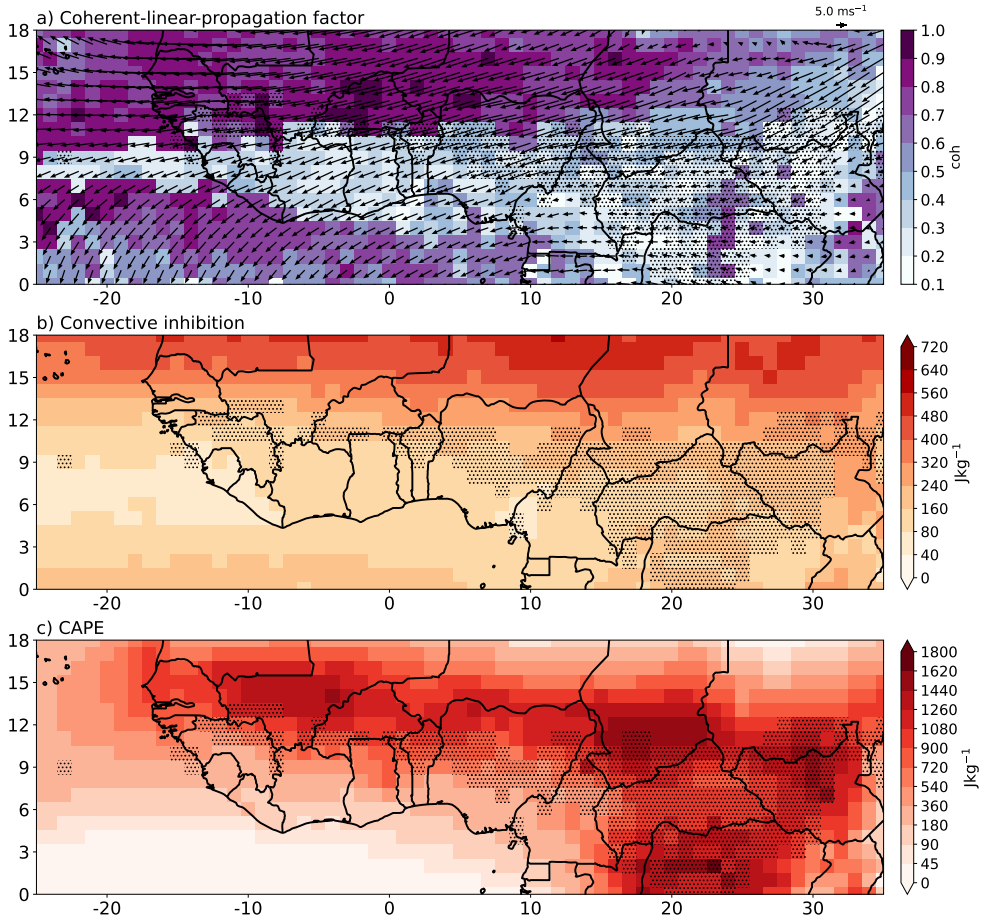



**Figure 6.2:** Diagnostic features: a) coherent-linear-propagation factor (coh) and vertical shear between 600 and 925 hPa (arrows), b) Convective INhibition (CIN), and (c) Convective Available Potential Energy (CAPE) for JAS 2001–2019 except 2016. Stippled grid points have a maximum CPA below 0.575 at all three lags. Reprinted from Rasheeda Satheesh et al. (2023). ©2023 The Authors, John Wiley & Sons Ltd.

for lags of two and three days, larger distances and more consistent westward propagation are found (Fig. 6.1e–f,h–i), suggesting that at these timescales propagating disturbances dominate over local effects. Vertical shear (Fig. 6.2a) in this region is mostly northeasterly and thus does not follow the largely easterly direction to the north of the rainbelt. Despite this propagation is predominantly westward, indicating that levels other than 600 and 925 hPa may play a role. Another characteristic of this region is low CIN (Fig. 6.2b), which slightly increases southward towards the equatorial cold tongue, and low CAPE (Fig. 6.2c). Given the absence of triggers common over land (diurnal heating, orography, coastal effects, etc.), this suggest that triggers from meso- or synoptic-scale circulation features may be necessary to release the weak convective instability or even to cause stratiform rain over the oceanic region. While Knippertz et al. (2017) describe slowly westward-moving, oceanic, synoptic-scale cyclonic systems during the DACCIWA campaign in 2016 and their effects on rainfall, there are no extant in-depth studies of this phenomenon.

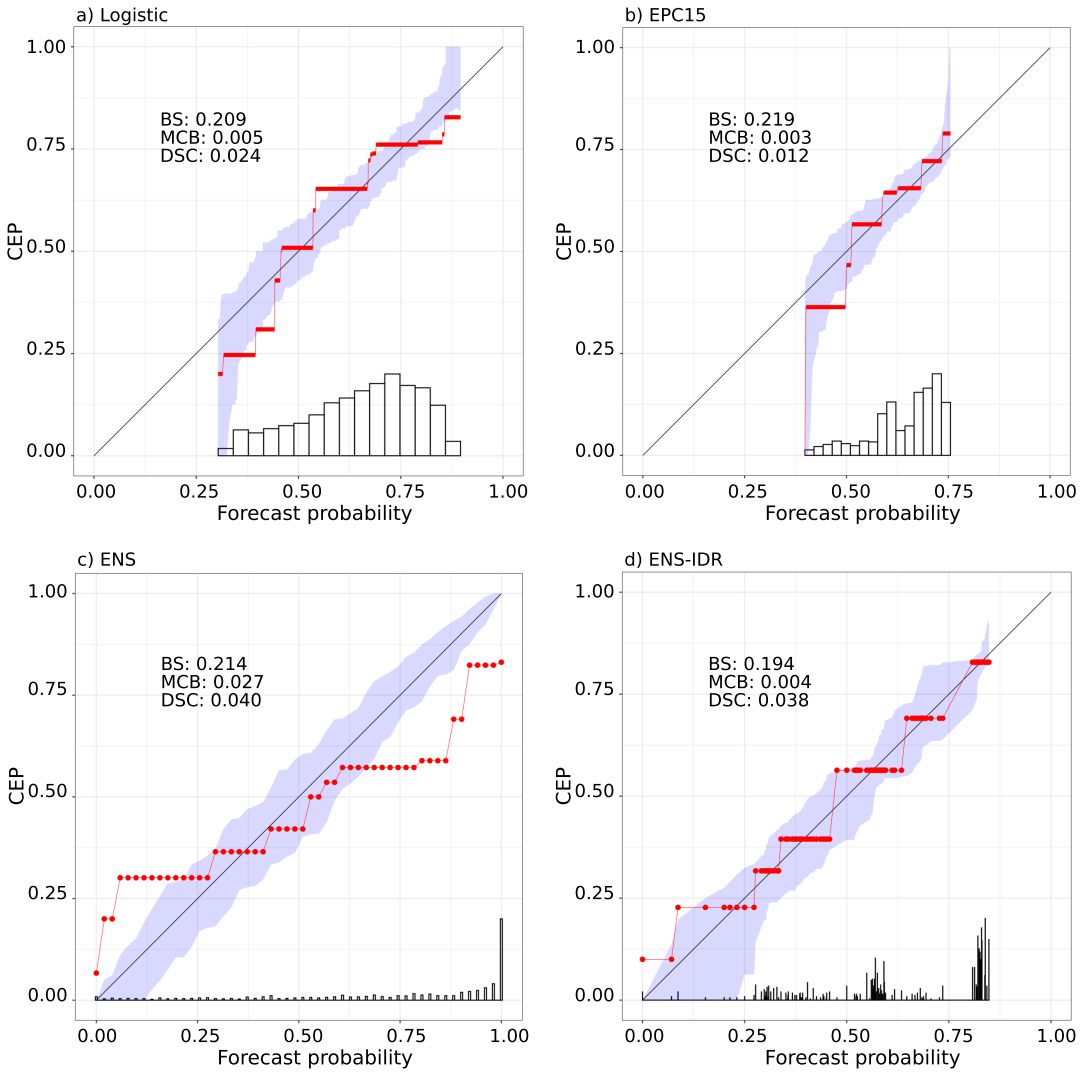



**Figure 6.3:** CORP reliability diagrams for a) Logistic, b) EPC15, c) ENS, and d) ENS-IDR probability of precipitation forecasts at the grid point closest to Niamey (13°N, 2°E) for JAS 2007–2019. The CORP reliability curves plot the conditional event probability (CEP) against the forecast value, along with 90% consistency bands under the assumption of perfect reliability. We also show the mean Brier score (BS) and its CORP miscalibration (MCB) and discrimination (DSC) components. The UNC component is 0.228 and only depends on climatology and not the forecast. Reprinted from Rasheeda Satheesh et al. (2023). ©2023 The Authors, John Wiley & Sons Ltd.

## 6.2 Forecast performance

As noted, our purely data-driven Logistic probability of precipitation forecast uses the CPA analysis to identify predictor variables from rainfall records. We now assess its predictive performance relative to the EPC15, ENS, and ENS-IDR forecasts.

For an initial illustration at the previously considered grid point closest to Niamey, Fig. 6.3 shows reliability diagrams in the recently proposed CORP form (Dimitriadis et al., 2021) for the Logistic forecast and the three competitors. The CORP approach assumes a monotonic, non-decreasing shape of the conditional event probability (CEP)

or reliability curve and yields an associated decomposition of the mean Brier score (BS) into miscalibration (MCB), discrimination (DSC), and uncertainty (UNC) components, namely,

$$\mathrm{BS} = \mathrm{MCB} - \mathrm{DSC} + \mathrm{UNC}, \tag{6.1}$$

while avoiding the shortcomings of the traditional binning and counting approach to reliability diagrams and score decomposition. For a comparison to the classical approach, see Fig. 6.4 and the detailed discussion in Dimitriadis et al. (2021). Supplementing the CORP reliability curves and confidence bands, which reflect the natural variability of the reliability curves under the assumption of perfect calibration, the inset histograms illustrate the marginal distribution of the forecast values. In the case of discrete probabilities (ENS, ENS-IDR) bar plots are used, and in the case of (nearly) continuously distributed forecast probabilities (Logistic, EPC15) density estimates are shown.

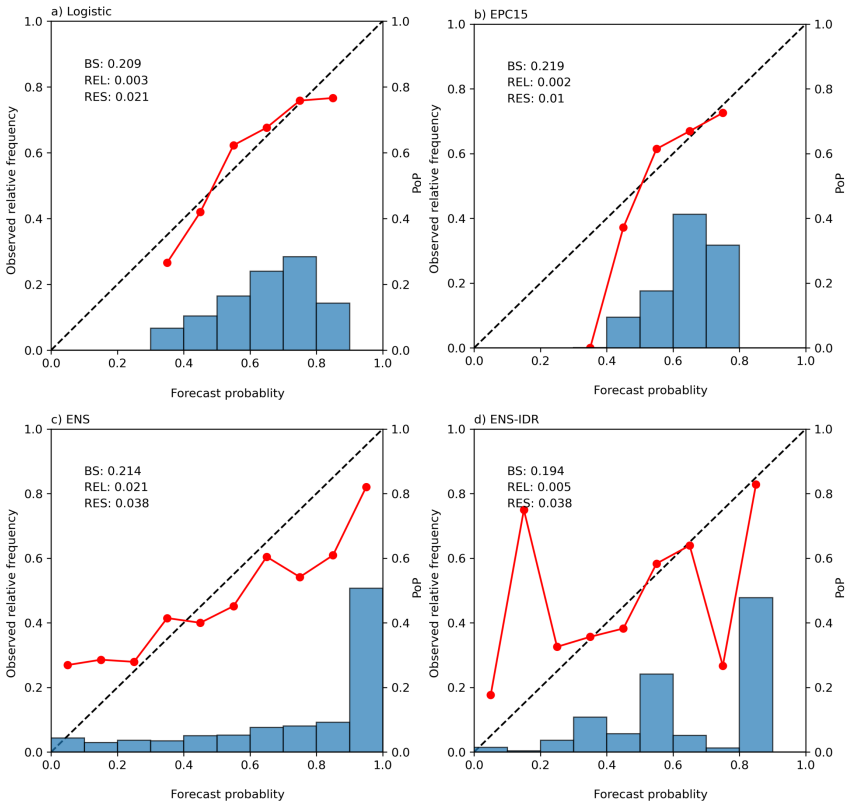



**Figure 6.4:** Traditional reliability diagrams for a) Logistic, b) EPC15, c) ENS, and d) ENS-IDR probability of precipitation forecasts at the grid point closest to Niamey (13°N, 2°E) for JAS 2007–2019. We also show the mean Brier score (BS) and its traditional reliability (REL) and resolution (RES) components, which assume the roles of the MCB and DSC components in Fig. 6.3. The UNC component is 0.228 and does not depend on the forecast nor on the type of score decomposition (CORP or traditional). Reprinted from Rasheeda Satheesh et al. (2023). ©2023 The Authors, John Wiley & Sons Ltd.

For the Niamey grid point, the Logistic forecast shows a mean Brier score of 0.209.[2] It is well calibrated (MCB = 0.005), shows moderate discrimination ability (DSC = 0.024) and forecast probability range $\sim 0.30 - 0.87$. In terms of the Brier score, it outperforms both the EPC15 and the ENS forecast but not the ENS-IDR forecast.

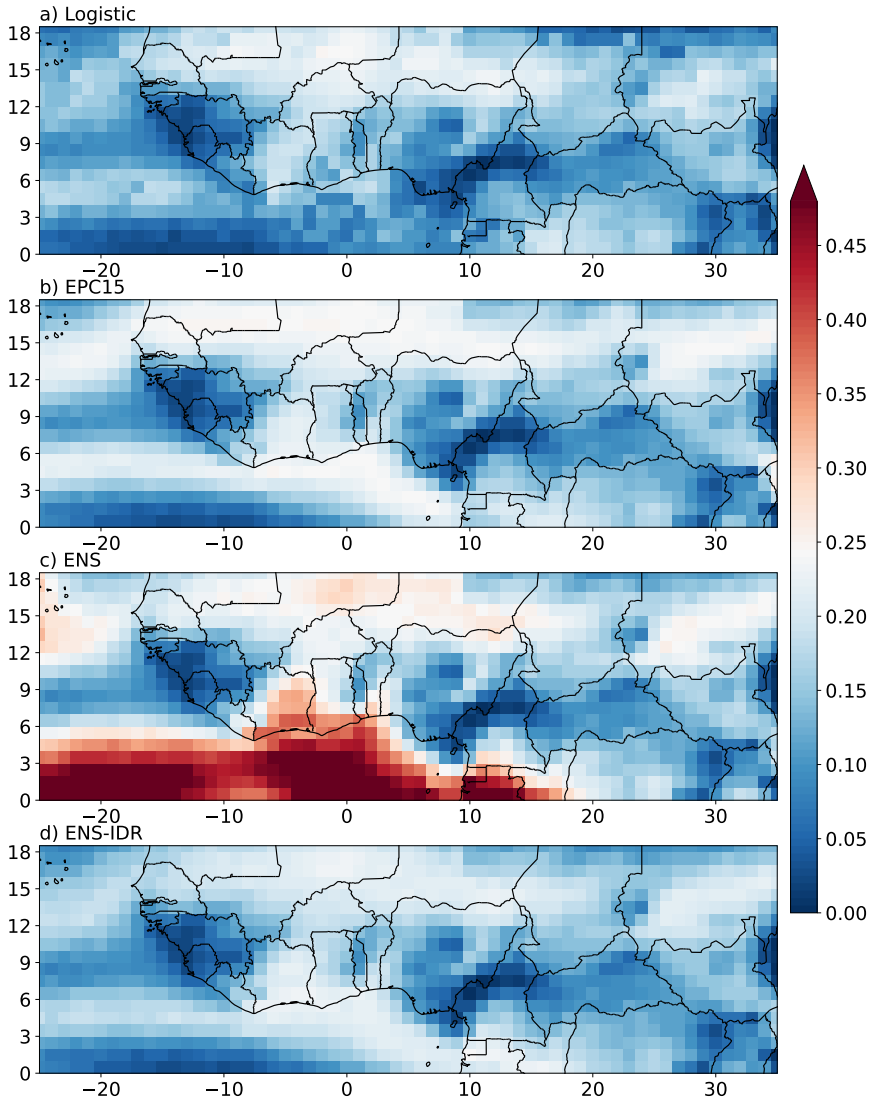



**Figure 6.5:** Mean Brier score (BS) of a) Logistic, b) EPC15, c) ENS, and d) ENS-IDR probability of precipitation forecasts for JAS 2007–2019. Reprinted from Rasheeda Satheesh et al. (2023). ©2023 The Authors, John Wiley & Sons Ltd.

The EPC15 forecast is even better calibrated (MCB = 0.003) but has poor discrimination ability (DSC = 0.012), with forecast probabilities ranging between 0.40 and 0.75 only. The low values in the marginal distribution are likely indicative of the monsoon onset, which is highly variable from year to year, while the high values indicate the peak of the monsoon season. In contrast, the ENS forecast is poorly calibrated (MCB = 0.027), but shows superior discrimination ability (DSC = 0.040), with forecast probabilities

[2]We repeated the analysis considering only lags up to two days, i.e., by setting $a_3 = 0$ in Eq. (5.3), which resulted in a slightly worse Brier Score (BS = 0.211), mainly due to a decrease in discrimination (DSC = 0.020) despite a slight improvement in calibration (MSC = 0.004).

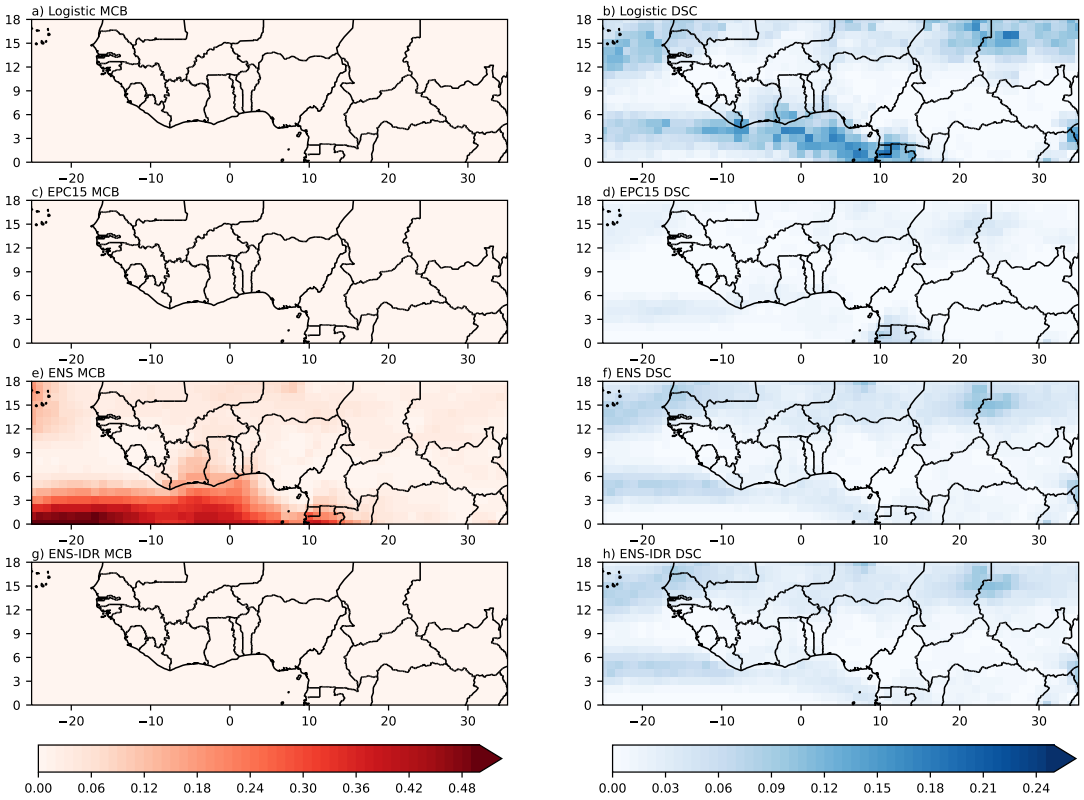



**Figure 6.6:** CORP miscalibration (MCB, a,c,e,g) and discrimination (DSC, b,d,f,h) components of the mean Brier score (BS) for (a–b) logistic, (c–d) EPC15, (e–f) ENS, and (g–h) ENS-IDR probability of precipitation forecasts for JAS 2007–2019. Reprinted from Rasheeda Satheesh et al. (2023). ©2023 The Authors, John Wiley & Sons Ltd.

attaining the full range from 0 to 1. However, when the rainfall probability of the ENS forecast is 1.00 (i.e., all 51 members indicating rain), which is surprisingly often (50.9%) the case, rainfall occurs with a conditional event probability (CEP) of 0.85 only. Finally, the postprocessed ENS-IDR forecast is well calibrated (MCB = 0.004), while retaining most of the discrimination ability from the ensemble (DSC = 0.038), leading to the best mean Brier score of 0.194, 0.015 better than Logistic. However, it is to be noted that the postprocessing reduces the range of forecast probabilities to $\sim 0.00 - 0.85$.

We repeat this comparison at every target grid point within the nested core area, to obtain the mean Brier score (BS, Fig. 6.5) and the associated MCB and DSC components (Fig. 6.6) from the CORP decomposition. The most striking result is that the ENS forecast suffers from poor calibration over much of the analysis region, and very poor calibration over the Gulf of Guinea and the equatorial Atlantic Ocean in particular (Fig. 6.6e). However, it needs to be mentioned that similar to Vogel et al. (2020) that pointed out that the uncertainties associated with TRMM estimates over the oceanic *deserts* may be responsible for the high miscalibration in ENS, uncertainties associated with the IMERG precipitation estimates (e.g., Bolvin et al., 2021) over the oceans may be partly responsible for the high MCB values of ENS in this study. The MCB values

of the postprocessed ENS-IDR forecasts resemble those of the Logistic and the EPC15 forecasts, with all three being well calibrated (Fig. 6.6a, c, and g). Not surprisingly, the climatological EPC15 forecast shows the least discrimination ability, as reflected by low DSC values (Fig. 6.6d). The spatial patterns of the DSC values for the Logistic forecast (Fig. 6.6b) resemble those for CPA at a lag of one day (Fig 6.1a) and peak in the off-equatorial Atlantic Ocean near the Cape Verde Islands, the eastern Sahel, the Gulf of Guinea, and the coasts of southern West Africa. DSC values for ENS and ENS-IDR (Fig. 6.6f and h) show some spatial agreement with Logistic but are generally much smoother. The good agreement between the two indicates that the postprocessing retains the discrimination ability of the dynamical model. Grid points within the African rainbelt generally show very low DSC values for all four forecasts considered, which is likely related to the high level of stochasticity discussed above.

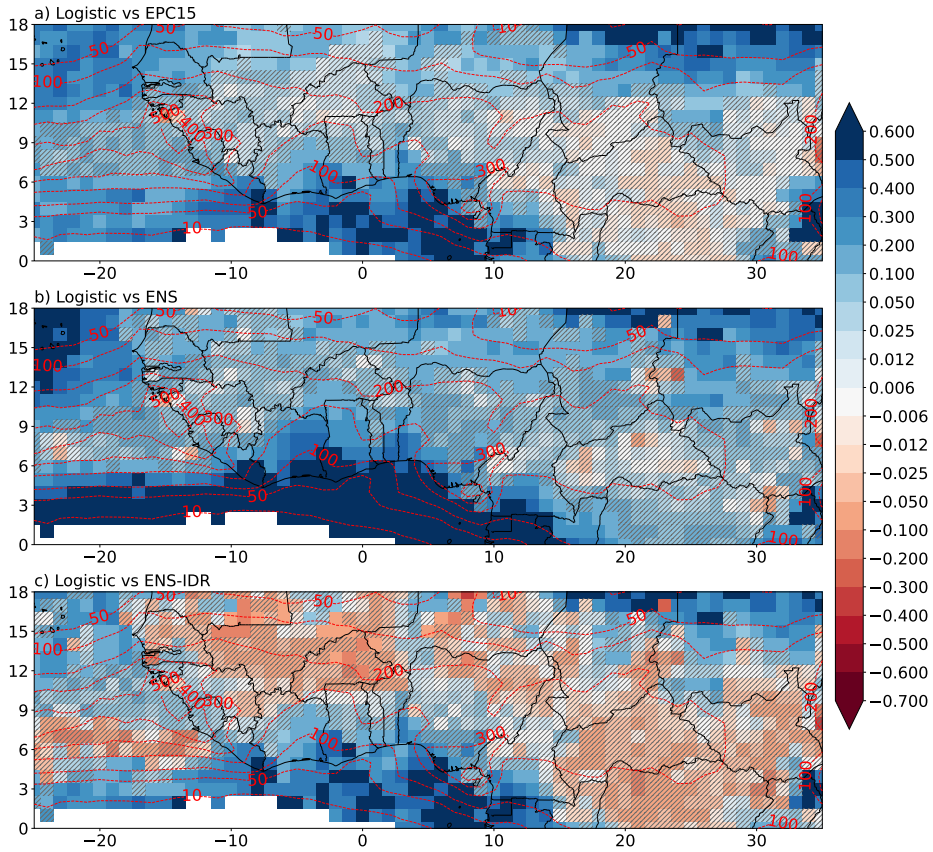



**Figure 6.7:** Brier skill score (BSS) for the Logistic forecast relative to the a) EPC15, b) ENS, and c) ENS-IDR forecasts for JAS 2007–2019. Hatching indicates grid points where the difference in the mean Brier score (BS) fails to be statistically significant at a level of 0.05, according to the Benjamini–Hochberg corrected Diebold–Mariano test for equal predictive ability. The contours show average monthly rainfall accumulation in mm. Grid points with less than 5mm of average monthly rainfall are left blank. Reprinted from Rasheeda Satheesh et al. (2023). ©2023 The Authors, John Wiley & Sons Ltd.

To facilitate a direct comparison between the Logistic forecast and the benchmark techniques, Fig. 6.7 shows the Brier skill score (BSS) of the Logistic forecast relative

to the EPC15, ENS, and ENS-IDR forecasts. Relative to EPC15 (Fig. 6.7a), the Logistic forecast shows better skill over most of our evaluation domain, particularly in regions of high rainfall gradients, where coherent (Fig 6.2a) propagating features create predictability through convective triggers or ambient conditions. Improvements over EPC15 in the rainbelt, however, are mostly small and statistically insignificant (hatched areas). In comparison to the ENS forecast (Fig. 6.7b), the Logistic forecast is superior practically everywhere, with some very high values over the ocean. Nonetheless, according to the Diebold–Mariano test for equal predictive ability, the differences between the Logistic and the EPC15 and ENS forecasts fail to be statistically significant over some parts of the climatological rainbelt.

The BSS of the Logistic forecast relative to the ENS-IDR forecast (Fig. 6.7c) shows an overall mixed result. The Logistic forecast performs significantly better over regions with high CPA, namely, the northeastern Sahel, the Gulf of Guinea, and the majority of the highly populated coastal regions along it (comprising Côte d'Ivoire, Ghana, Togo, Benin, Nigeria, and Cameroon), the oceanic regions surrounding the Cape Verde islands near the coast of Senegal, and nearby coastal regions, due to superior discrimination ability (Fig. 6.6b). The ENS-IDR forecast performs better over a narrow strip of grid points over the Atlantic Ocean off the coast of Sierra Leone along the southern boundary of the rainbelt, in the western Sahel, in the Central African Republic, and in the Congo basin. However, the associated differences in the mean Brier score fail to be statistically significant. This confirms that even with sophisticated postprocessing, a forecast with a dynamical model can hardly outperform a relatively simple data-driven one.

Interestingly, the findings in this section remain largely unchanged if the Logistic forecast is based on predictor variables at a temporal lag of one day only, barring some deterioration of the predictive performance over the western Sahel (Fig. 6.8). This somewhat surprising result indicates that the bulk of the useful information is contained in yesterday's rainfall, be it a local storm that persists from one day to the next (as often the case over the ocean) or an MCS or AEW that propagates towards the point of interest (as often the case over the Sahel).

## 6.3 Highlights of pre-monsoon and transition seasons

As previously stated, this section will repeat the previous analyses for the seasons of May-June, March-April, and October-November, and notable findings will be discussed in detail.

The analysis of the maximum CPA for MJ (Fig. 6.9a–c) displays patterns akin to JAS, albeit with a minor southward shift in line with the southward shift of the rainbelt. As a result, a larger area of northern tropical Africa receives very low rainfall (less than 10 mm per month), and the regions along the Gulf of Guinea receive abundant rainfall (more than 300 mm per month over the coast and $50-100$ mm per month along the equator). The overall CPA patterns are comparable to JAS, with high CPA along rainfall gradients and low CPA in regions coinciding with the climatological rainbelt.

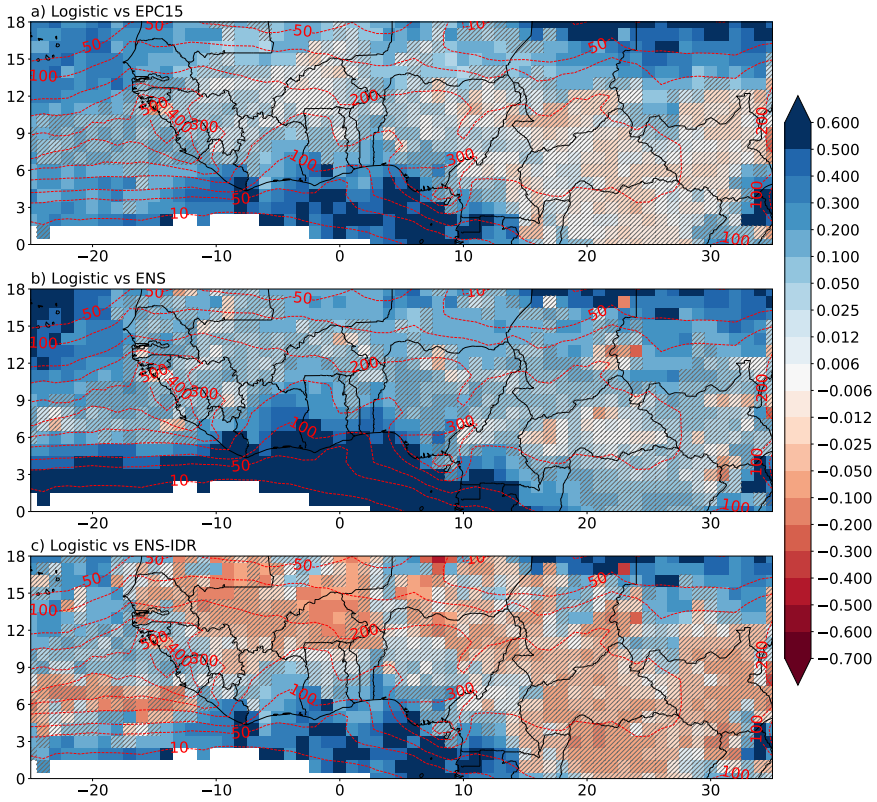



**Figure 6.8:** Same as Fig. 6.7 but for a logistic model that uses spatio-temporal predictors at a lag of one day only, by setting $a_2 = a_3 = 0$ in Eq. (5.3). Reprinted from Rasheeda Satheesh et al. (2023). ©2023 The Authors, John Wiley & Sons Ltd.

The distances on the one-day lag (Fig. 6.9d) are largely around 400–500 km, indicating local influences over most of the analysis region except over the rainbelt region (mainly central Sahel and central Africa). At higher lags (Fig. 6.9e–f), clearer patterns emerge across the majority of the analysis area with the exception of the oceanic region near the Senegal coast and the far northeastern Sahel. The directions shown in Fig. 6.9g–i indicate predominantly westward propagation across most of the grid points in the Sahel region, with dominant eastward propagation observed over central Africa and partially over the oceanic region adjacent to Sierra Leone's coast. The westward propagation of all three waves, covering approximately 500-800 km per day (around 6-9 m/s), may suggest weak AEWs as AEJ is relatively weaker in May-June (Fink et al., 2017). The propagation observed over the Congo basin towards the east may suggest the presence of Kelvin waves, as prior research has linked synoptic-scale rainfall in the Congo basin to Kelvin waves in the boreal Spring–Summer seasons (e.g., Mounier et al., 2007; Wang and Fu, 2007; Nguyen and Duvel, 2008; Sinclaire et al., 2015).

A significant deviation from these overall patterns is the northward excursion of higher rainfall and correspondingly lower (than surrounding) CPAs observed over central Mali on all lags. The greater distances observed throughout this region, combined with erratic directions on higher lags, suggest a lack of coherent propagation, which is generally observed in stochastic environments such as the rainbelt. Vizy and Cook (2009, 2014)

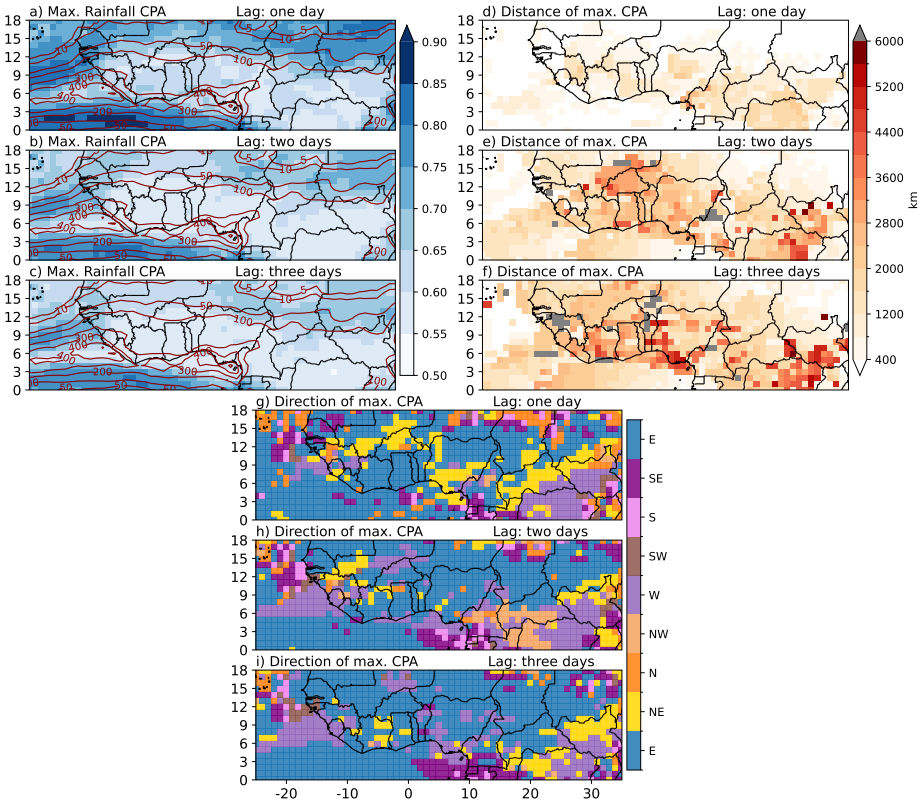



**Figure 6.9:** Same as Fig. 6.1 but for the MJ season.

explore how cold air surges from the Mediterranean to northern Africa contribute to monsoon breaks in the eastern Sahel and amplify synoptic-scale convective activity in the western Sahel. Such a mechanism may be a possible explanation for the northward shift of the rain belt observed only over Mali.

The coherence patterns observed (see Fig. A1a) are comparable to the patterns of maximum CPAs shown in Figs. 6.9a–c. Due to the absence of robust AEWs in MJ, the distinction between regions of high and low coh is not as apparent as in JAS. However, similar patterns can still be discerned, notably high coherence over the outskirts of the rainbelt and low coherence in the stochastic regions which coincide with the climatological rainbelt. The convective inhibition (CIN) in this period is mostly similar to that of JAS, but there is a significant shift towards the South for high CIN values (above 320 $Jkg^{-1}$). This shift is aligned with the southward movement of the rainbelt. The CAPE patterns, which are not depicted, bear similarity to those of JAS but with a southern shift. Except for northern Mali, the far northern regions exhibit extremely low CAPE (less than 45 $Jkg^{-1}$), which is in accordance with the rainfall patterns observed (as seen in the dark red isohyets in Figs. 6.9a–c).

Most of the patterns seen in the BSS (Fig. 6.10) of the Logistic forecast against the benchmarks are similar to those seen in the JAS season. The Logistic forecast outperforms the EPC15 forecast (Fig. 6.10a) across the majority of the analysis region. Over the rain belt region, particularly across the Congo basin and parts of the Sahel, the

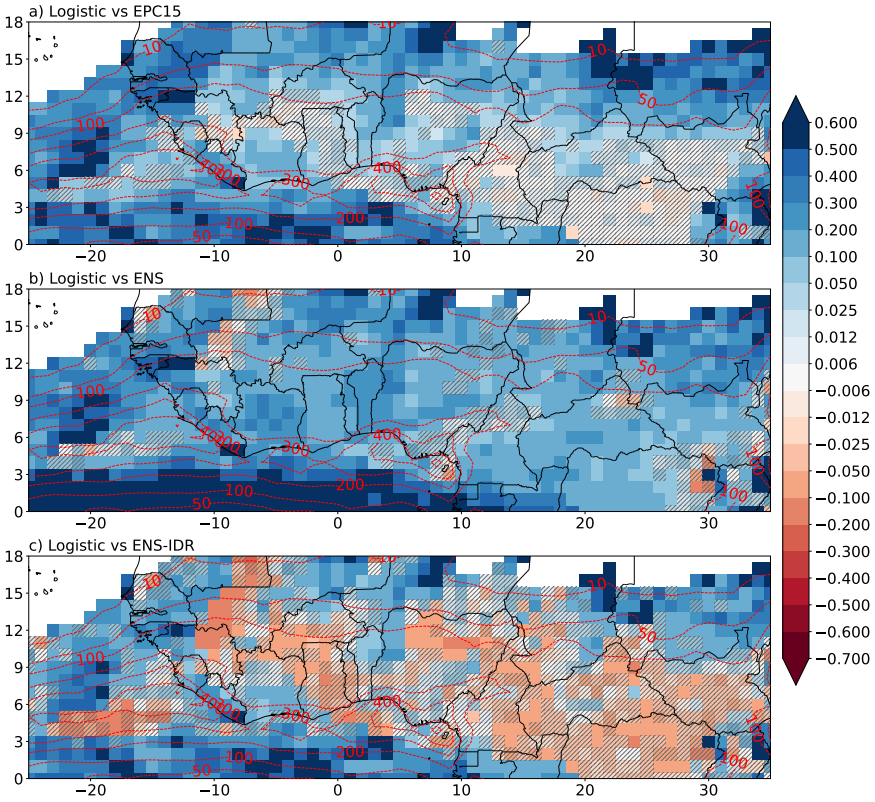



**Figure 6.10:** Same as Fig. 6.7 but for the MJ season.

differences between the Logistic and EPC15 forecasts fail to reach statistical significance according to the Diebold-Mariano test (shown as hatched grids) (Diebold and Mariano, 2002). In comparison with the ENS forecast (Fig. 6.10b), the Logistic forecast demonstrates significant superiority nearly everywhere, except for some grid points near the Cameroon highlands and some over eastern Congo and Uganda, where the differences are not significant. Similar to the JAS period, the Brier Skill Score (BSS) of the Logistic forecast compared to the ENS-IDR forecast (Fig. 6.10c) presents a mixed result, but shifted southwards in line with the shift of the rain belt. While the Logistic forecast outperforms the ENS-IDR forecast throughout most regions outside of the rainbelt, the latter performs better in the rainbelt area. It should be noted, however, that apart from some grid points in the western Sahel (mainly northern Guinea and southern Mali), northern Nigeria and small pockets over central Africa, the differences between the logistic forecast and the ENS-IDR forecast are not significant.

The logistical forecast outperforms the ENS and ENS-IDR forecasts over the Gulf of Guinea, with BSS values exceeding 0.6 at certain grid points. The region experiences significant precipitation, around 100-200 mm per month, as indicated by the red isohyets in Fig. 6.10, but lacks common triggers seen over land, such as orography, requiring a strong driver for rainfall. The cyclonic vortices described in Knippertz et al. (2017) can serve as the necessary triggers to facilitate rainfall by overcoming the low CIN observed over this region.

Figs. 6.11a-c illustrate the maximum CPAs achieved over the analysis region during the MA season. The CPA patterns indicate a shift towards the equator compared to both the MJ and JAS seasons, which is in line with the location of the rain belt being much closer to the equator. Due to the rain belt's equatorward relocation, a significant portion of the analysis region (north of $\sim 13^\circ$ N over land and $\sim 7^\circ$ N over sea) experiences little to no rainfall (less than 5 mm per month). As noted in the previously discussed chapters, regions with high rainfall gradients tend to have high CPAs, whilst CPAs in regions coinciding with the rain belt (south of $\sim 8^\circ$ N over land and south of $\sim 6^\circ$ N over the ocean) are typically lower. However, an exception to these trends is apparent over the Atlantic, exhibiting significant CPAs on all three time lags (reaching up to 0.70, even on a three-day lag).

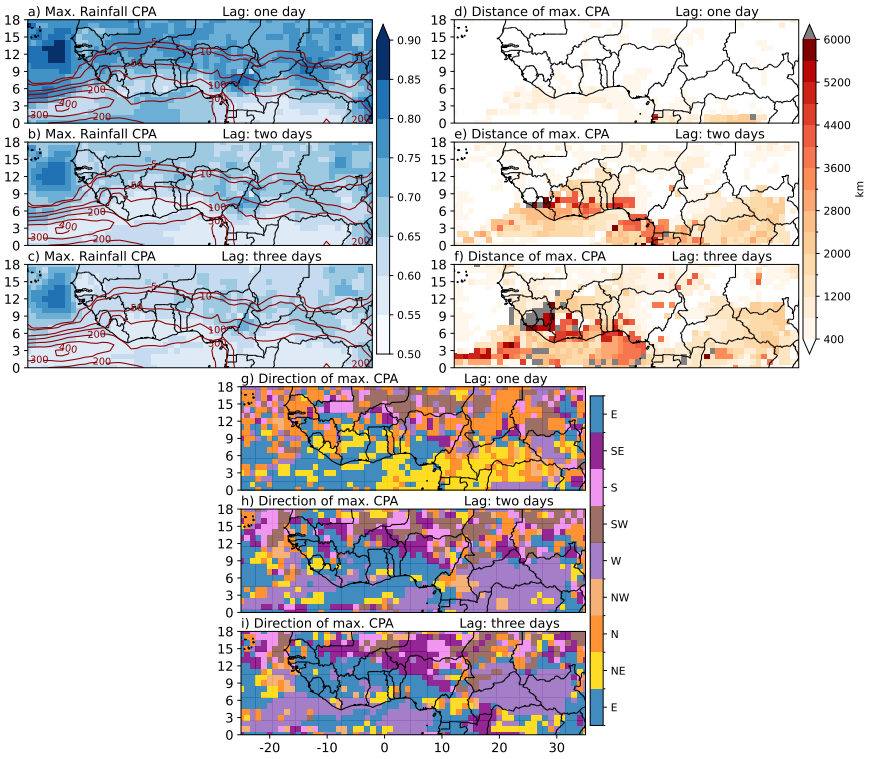



**Figure 6.11:** Same as Fig. 6.1 but for the MA season.

Distances and directions of the maximum CPAs in MA (Fig. 6.11d-f and Fig. 6.11g-i, respectively) indicate the prevalence of local interactions over most of the analyzed area with a one-day lag. The distances are mostly around 400 km or shorter. The directions do not exhibit any discernible pattern, although southward propagation is the most frequent, along with occasional pockets of northwestward and northward propagation. Interactions remain local in the regions of high CPAs, even at longer lags, while directional patterns exhibit greater clarity, with a prevailing north-northwestward propagation. Although patterns seen on a one-day lag may be stochastic, those on longer lags may be attributed to the tropical plume-like structures arising from tropical-extratropical interactions (Knippertz and Martin, 2005). As described by Knippertz and

Martin (2005), these weather systems are also (rarely) capable of inducing substantial precipitation across northern tropical Africa, an area that is typically arid during the boreal winter–spring seasons.

Higher lags show larger distances of maximum CPAs over the rainbelt region, with eastward being the most dominant direction of propagation mixed with smaller regions displaying westward propagation. The eastward propagation can at least partially be attributed to convectively coupled Kelvin waves, as noted in some previous studies (e.g., Wang and Fu, 2007; Sinclaire et al., 2015). However, it is probable that the observed westward propagation is artefactual due to the fact that the distances of maximum CPAs on higher lags surpass 6000km, and CPAs over such large distances are likely not physical.

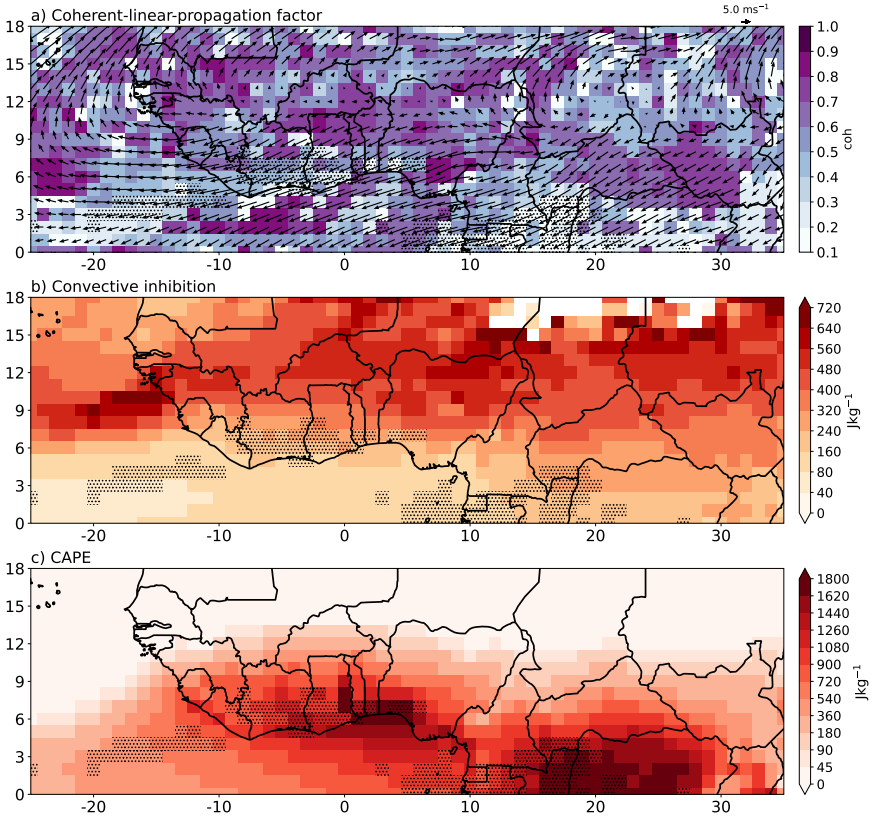



**Figure 6.12:** Same as Fig. 6.1 but for the MA season.

The coherence values in MA (Fig. 6.12a) lack clear patterns, unlike those observed during MJ and JAS seasons. Pockets of high coh are seen in central Africa (Central African Republic and parts of South Sudan), in the Sahel (northern Nigeria, Benin, Togo and Ghana) and over the Atlantic Ocean, but all these are likely caused due to short-ranged interactions as shown by small distances the largest of which is over central Africa (around 2000 km in three days). A combination of very high CIN (Fig. 6.12b) and low CAPE (Fig. 6.12c) north of $\sim 9°$ N indicate unfavourable conditions for rainfall, which is in line with the location of the rainbelt. Nonetheless, elevated CAPE values (over 1800 Jkg$^{-1}$) and moderate CIN offer highly favourable conditions for convection in

the Congo basin and the coast of Nigeria. This may explain the inconsistent propagation seen over the three lags at these locations, as such an environment is very stochastic. Additionally, the reduced coh values observed at these locations reinforce this argument.

The Logistic forecast demonstrates superior performance compared to both the EPC15 and ENS forecasts across the majority of the analysis area with regard to the BSS (see Fig. A2). At a few grid points in the Congo Basin and over Gabon, however, the differences between the Logistic and EPC15 forecasts are not significant. Similar to the MJ and JAS periods, the BSS of the Logistic forecast in relation to the ENS-IDR forecast reveals mixed outcomes. The Logistic forecast reveals greater proficiency in areas which coincide with the fringes of the rainbelt and, interestingly, over sections of the northern DRC, which is situated in the rainbelt. In contrast, the ENS-IDR forecast shows only slight (not significant) improvements over some regions in the rainbelt (mainly parts of the Congo basin and Cameroon).

The highest values of maximum CPAs achieved during the ON season over the analysis region (illustrated in Figs. 6.13a–c) showcase distinct patterns compared to the other seasons. It exhibits patterns that are almost anti-symmetric to those found in MA. The highest values of CPAs are seen on the one-day lag in the Sahel, with CPAs exceeding 0.90 observed over several grid points in central Nigeria, the northern Central African Republic and Sudan. In contrast to other seasons, high CPAs sometimes appear in areas with low rainfall gradients. For instance, a number of grid points in Nigeria and northern Cameroon receive rainfall of over 100 mm per month yet exhibit CPAs of 0.8 or higher. CPAs of up to 0.75 are commonly seen across the Sahel, particularly in a band stretching between 16° N and 6° N, even on higher lags. The area that aligns with the centre of the rainfall belt lies to the south of this band of high CPAs and contains grid points with CPAs as low as 0.5 at higher lags.

The distances of maximum CPAs (illustrated in Figs. 6.13d–f) demonstrate that local interactions (400–800 km) are the primary source of high CPAs on the one-day lag. Only the grid boxes in the Congo Basin show slightly larger distances (around 1200 km). The distances increase gradually on higher lags. The distances observed on the three-day lag denote propagation of approximately 400 to 700 km per day (about 4.5–8 $ms^{-1}$). The directions indicated in Figure 6.13g suggest that, for the most part, there is westward, northwestward, and southwestward propagation over the Sahel and oceanic region, with eastward propagation over the Congo basin. The directions on higher lags (Fig. 6.13h–i) reveal a pattern of well-organized westward movement across extensive areas of the Sahel, with occasional instances of northwestward propagation over the northwestern Sahel and southwestward propagation over the eastern Sahel reaching as far south as the Congo Basin.

The coherent distances and directions of maximum CPA described above translate to very high coh (mostly 0.6–0.8) values (Fig. 6.14a) over almost the entire Sahel region in ON. Moreover, a combination of moderate to high values of CIN (shown in

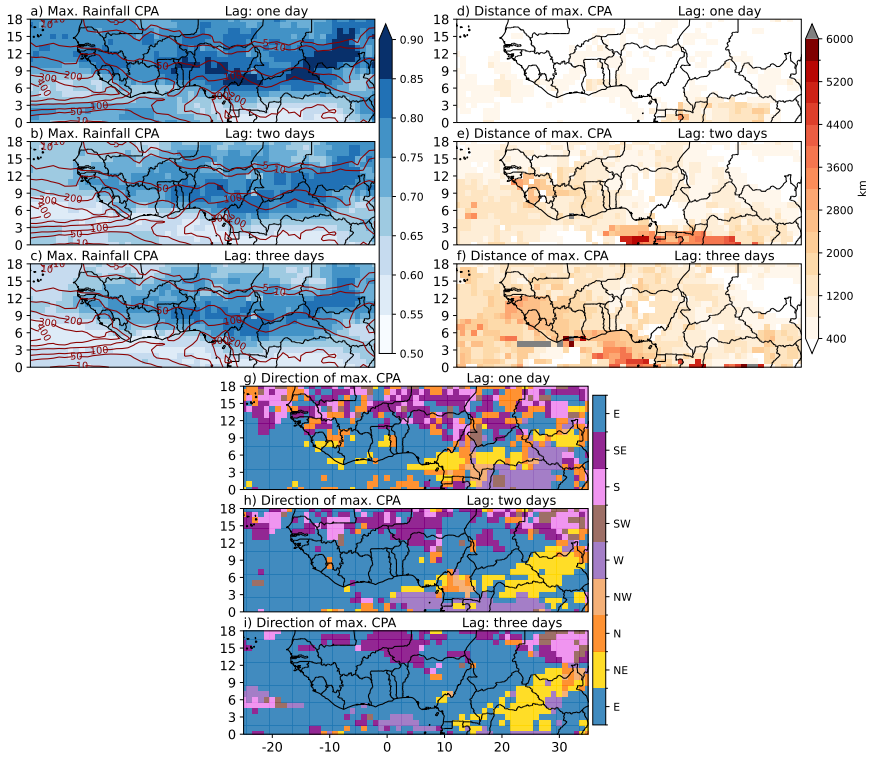



**Figure 6.13:** Same as Fig. 6.1 but for the ON season.

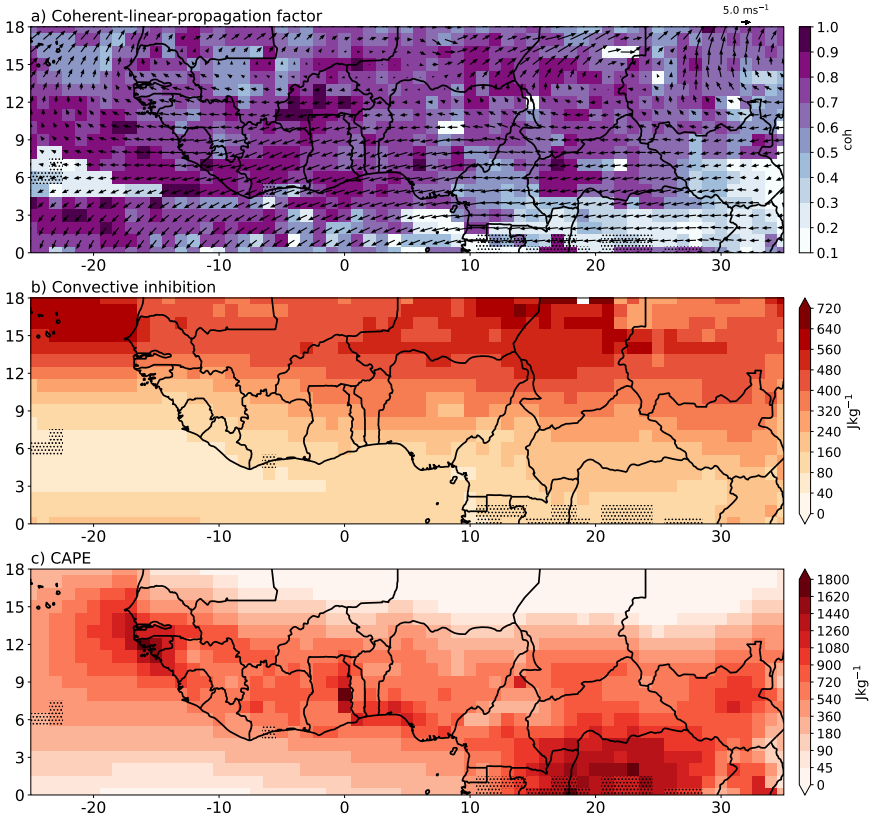



**Figure 6.14:** Same as Fig. 6.1 but for the ON season.

Fig. 6.14b) with fairly high values of CAPE (shown in Fig. 6.14c) across the Sahel region indicates that a coherent driver such as the AEWs can induce and maintain high levels of precipitation.The shear values illustrated in Fig. 6.14a (vectors) are weaker than those presented in JAS (Fig. 6.2a), suggesting that the MCSs generated are weaker in comparison, further highlighting the need for a coherent forcing like the AEWs. The areas that exhibit low coherence are located in the Congo basin, Gabon, and the neighbouring oceanic region in the Gulf of Guinea, which are situated at the center of the rain belt, thereby making them highly stochastic in nature. This is further supported by the combination of the high CAPE and low CIN values seen over these regions. A smaller pocket of low coh can also be seen in the western oceanic region off the coast of Liberia and Sierra Leone, although the exact reason for this is not clear.

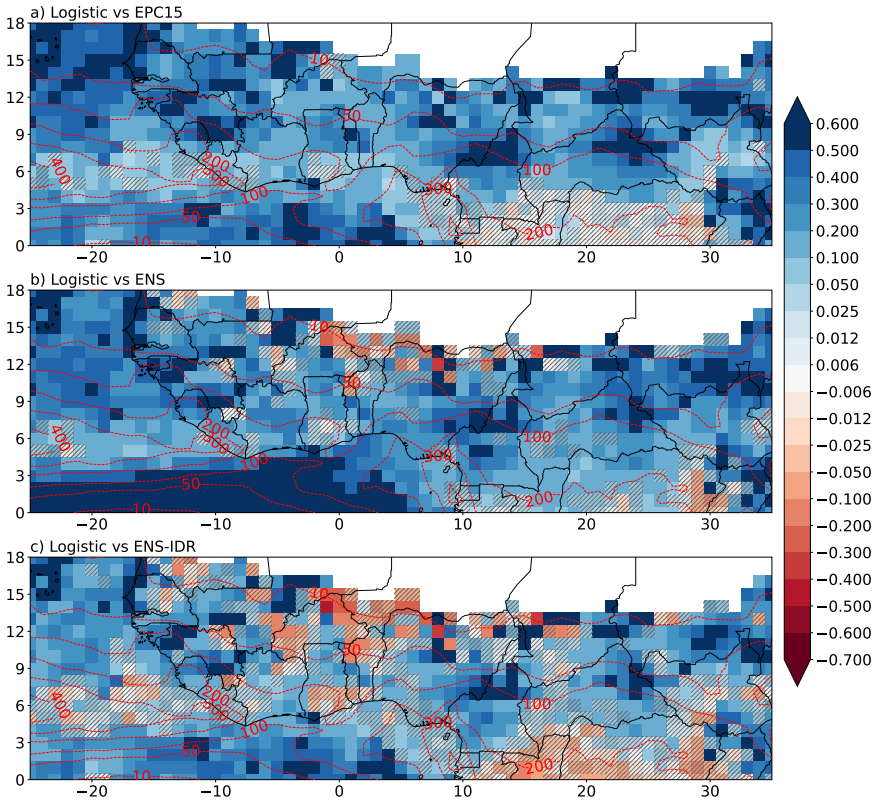



**Figure 6.15:** Same as Fig. 6.7 but for the ON season.

Figure 6.15 presents the Brier Skill Score (BSS) of the Logistic forecast in comparison to the benchmark forecasts. It is evident that the Logistic forecast outperforms the EPC15 forecast (Figure 6.15a) across most of the analysis area, with BSS values surpassing 0.6 at several grid points. While the BSS values are generally concentrated within the 'band' of high CPA values presented in Figs. 6.13a–c, a number of grid points located over the Gulf of Guinea and the offshore area adjacent to Senegal display BSS values of up to 0.6. At the grid points situated predominantly in the areas aligned with the rainbelt (the Congo Basin, Gabon, Equatorial Guinea and the adjacent oceanic region), the differences between the Logistic and EPC15 forecasts are not significant according to the Diebold-Mariano test (hatching in Fig. 6.15a).

The Logistic forecast demonstrates superior performance to the ENS forecast (Fig. 6.15b) across most areas, except for specific grid points located within the western Sahel (primarily southern Mali, southern Niger, northern Nigeria, and certain parts of Togo) and the Congo Basin. Despite this, the slight increase in the ENS forecast's skill is not statistically significant, as shown by the hatching. The performance of the Logistic forecast over the Gulf of Guinea is especially noteworthy, with BSS values above or equal to 0.5.

As observed in other seasons, the Brier Skill Score (BSS) comparison of the Logistic forecast to the ENS-IDR forecast displays mixed results. The ENS-IDR forecast surpasses the Logistic forecast in several grid points located in the Sahel region (particularly southern Mali, southern Niger, northern Nigeria and sections of Togo) and the Congo basin. Nevertheless, the differences between the Logistic and ENS-IDR forecasts are not significant. The Logistic forecast exhibits considerably better performance than the ENS-IDR forecast almost everywhere else.

## 6.4 Summary

Operational forecasts of rainfall in tropical Africa and the adjacent Atlantic Ocean have low skill (Vogel et al., 2018), but purely data-driven statistical models that exploit spatio-temporal correlations show promise (Vogel et al., 2021). To more systematically explore the potential of the data-driven approach, we have used the recently developed CPA in a spatio-temporal analysis of 24-hour accumulated GPM-IMERG rainfall from 2001—2019 for the West African summer monsoon season, spanning up to lags of three days temporally and covering up to several thousand km spatially. The CPA patterns found can be attributed predominantly to synoptic-scale drivers causing characteristic time-space behaviour of precipitation anomalies. To substantiate this, we have introduced a novel coherent-linear-propagation factor (coh) that quantifies the extent to which the locations of lagged maximum CPA reflect a propagation with constant phase speed and direction. Large values of coh combined with substantial CPA over the three days indicate physically interpretable and statistically stable relationships and, thus, the potential for high statistical predictability. Such patterns are usually associated with high values of CAPE, CIN and shear, which necessitates the presence of relatively strong triggers like AEWs and MCSs to generate rainfall. Land regions located within the climatological rainbelt show very low values of CPA and coh, indicating high stochasticity, while oceanic regions located within the rainbelt show sizeable CPA and moderate coh. High values of coh along the northern and southern fringes of the rainbelt over land (except in the ON season) reflect the presence of a coherent synoptic forcing that triggers and sustains precipitation over the three days.

When comparing the findings from the pre-monsoon period (MJ) and the transition seasons (MA and ON) with those from JAS, it is clear that weak AEWs account for the majority of the coherent features observed in both MJ and ON. Nevertheless, compared

to JAS, the distinction between regions of high and low coh is less clear, highlighting the influence of the strong AEWs on the daily rainfall occurrence during the summer monsoon season over tropical Africa. The discovery of almost anti-symmetric maximum CPA patterns in ON compared to MA, despite both being transition seasons, is an intriguing and unexpected finding. A potential reason for such patterns could be the presence of weak AEWs modulating the squall line-dominated precipitation over the Sahel in ON, which is relatively drier in MA.

It has been observed that cyclonic vortices propagating towards the west (Knippertz et al., 2017), which are the probable cause of the coherent westward propagation in the Gulf of Guinea in JAS, may be active from May to November. During the March-June period, which includes the MA and MJ seasons, tropical-extratropical interactions gain greater significance in the northern part of the analysis region. More precisely, it is likely that tropical plume-like events are responsible for a small yet substantial proportion of rainfall incidents in the northern regions of MA. Cold air surges are likely responsible for the northward movement of the rainbelt across the western Sahel and the decrease in rainfall across the eastern Sahel during MJ. Both MA and MJ demonstrate greater eastward propagation of maximum CPAs over central Africa and the Gulf of Guinea regions. Kelvin waves may be a contributing factor to these patterns, as they are known to modulate MCS triggers over central Africa (Laing et al., 2011).

To leverage these findings, we have carried out a systematic comparison of the predictive performance of a purely data-driven Logistic probability of precipitation forecast that employs predictors from the identified locations of maximum CPAs with three benchmark forecasts, namely, the climatological EPC15 forecast, the ENS forecast, which is heavily miscalibrated, and ENS-IDR. All forecasts struggle over regions where the precipitation is dominated by highly stochastic processes – in particular, in the African rainbelt in JAS – due to the lack of coherent synoptic-scale signals. In these regions, the differences between the mean Brier skill scores of the Logistic, the EPC15, and the ENS-IDR forecasts are small and fail to be statistically significant. However, along the fringes of the rainbelt over land and the Gulf of Guinea, where stochasticity is much less pronounced, statistical approaches benefit from dominant forcing that triggers and sustains precipitation events, such as AEW and the westward propagating cyclonic vortices described in Knippertz et al. (2017), well in line with earlier results by Vogel et al. (2021).

# 7. Relationship between NWP model forecast skill and tropical wave phases

Based on the previous chapter's findings, it becomes evident that utilising the predictive information contained in coherently propagating precipitation patterns can generate entirely data-driven forecasts of daily rainfall over large regions within tropical Africa. The aforementioned forecasts have been proven to surpass the skill of current NWP model-based precipitation forecasts, primarily in regions corresponding to the fringes of the climatological rainbelt. Even in the rainbelt region, over which rainfall is notoriously unpredictable, the statistical forecast shows skill marginally better than that of the NWP model-based forecast. As described in Chapter 2, convectively coupled tropical waves represent a significant proportion of the coherently propagating rainfall patterns in the tropics. It has been explained that the distinct phases of these waves, whether dry or wet, dictate the variability of precipitation across different regions in tropical Africa. Consequently, it is justifiable to explore their impact on the modulation of rainfall anomalies and the skill of NWP model-based forecasts in tropical Africa, both in terms of precipitation occurrence (at a threshold of 0.2 mm per day) and amounts. To this end, we structure this chapter as follows:

As explained in Chapter 5, a gradient boosting tool was developed to identify tropical waves-based predictors for training statistical models in Chapter 8. Using this tool, specific regions with the most significant rainfall modulation capacity for each of the seven TWs are identified by repeating the analysis described in Section 5.2.4 over the entire analysis domain. The dates corresponding to various phases of the waves are identified as described in Section 5.2.3 for performing a composite analysis of the daily rainfall anomalies modulation. Next, the area-averaged skill scores (BSS for rainfall occurrence and CRPSS for rainfall amount) are computed for the raw ensemble forecasts, relative to the EPC15 ensemble forecast (reference forecast), for every wave phase from $1-8$ in the regions identified. Finally, the findings of this chapter are summarized in the Summary section.

## 7.1 Identification of regions of highest wave activities

The analysis described in Section 5.2.4 is repeated over the entire domain to obtain an overview of the top predictor at each target grid point for all summer seasons, as shown

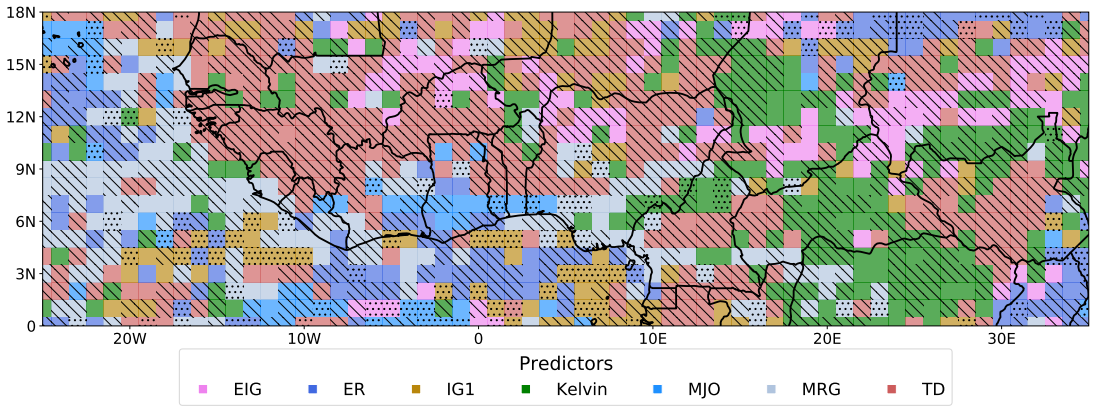



**Figure 7.1:** Most frequent top predictors, derived from the filtered daily rainfall anomaly, are presented for every target grid point across summer seasons (JAS) from 2007-2016 over the analysis domain. The hatched grid points indicate where the dominant predictor is derived from the grid point downstream to the target (in the direction of the respective wave's propagation), while stippled grid points indicate where the dominant predictor is derived from the upstream grid point. Dominant predictors at unmarked (no hatching/stippling) grid points originate from the same target grid point.

in Figure 7.1. Although the map looks quite noisy at first sight, a closer look reveals some discernible patterns and the following deductions can be drawn: 1) Approximately 61% (see table 7.1) top predictors come from the grid point downstream of the target grid point. Roughly 29% of the dominant predictors come from the target grid point itself, while $\sim$ 10% are selected from the upstream grid point; 2) TD and Kelvin waves are the two most dominant wave modes in the analysis region accounting for $\sim$ 48% (see table 7.1) of all grid points. Since Kelvin waves are wave modes that are divergent, it is likely that they couple with convection and follow the rainbelt to off-equatorial latitudes. This can account for the Kelvin wave activity spotted over the off-equatorial regions despite the theoretical Kelvin waves (or dry) being equatorially trapped. Nevertheless, as previously explained, the Kelvin wave activity at higher latitudes (northward of $\sim$ 15° N) is unlikely to be true Kelvin waves, but rather signals sharing the same space in the wavenumber-frequency power spectrum. On a similar note, it is worth highlighting that while the TDs over the Sahel and regions further north are representative of AEWs in the summer season, those found in the vicinity of the equator are unlikely to be equivalent to AEWs since they are far away from the effects of the African easterly jet (AEJ);

3) Not surprisingly, TDs that act as proxies for AEWs in the summer dominate as the top predictor (almost all of them downstream) over large regions of the Sahel from west to east, with few grid points of Kelvin and EIG waves. Some studies based on the Global Forecast System (GFS) models (e.g., Ventrice and Thorncroft, 2013; Lawton et al., 2022), have explored the link between AEWs and Kelvin waves, wherein one influences the genesis and propagation of the other via transfer of eddy kinetic energy (EKE). The mixed signals of AEWs and Kelvin waves as top predictors over the Sahel may be a

| Wave | Target | Downstream | Upstream | Total | % |
|---|---|---|---|---|---|
| EIG | 38 | 62 | 6 | 106 | **9.1** |
| ER | 47 | 68 | 30 | 145 | **12.5** |
| IG1 | 30 | 56 | 32 | 118 | **10.2** |
| Kelvin | 92 | 122 | 8 | 222 | **19.2** |
| MJO | 20 | 22 | 16 | 58 | **5** |
| MRG | 60 | 91 | 24 | 175 | **15.1** |
| TD | 45 | 285 | 5 | 335 | **28.9** |
| | **28.6** | **60.9** | **10.4** | | % |

**Table 7.1:** Total number of grid points where each tropical wave feature is identified as the primary predictor in the analysis domain.

result of such interactions. The situation is much less chaotic over the coastal regions to the west, however, and TD is clearly the dominant wave mode over this region. However, over the oceanic regions over the western coast, MRG waves are seen to be more active than AEWs; 4) Interestingly, in the analysis domain, slow-moving, large-scale waves such as the ER and the MJO together appear as dominant predictors at approximately 18% of the grid points. However, upon closer inspection, it becomes evident that these waves always occur alongside faster waves, such as TDs (not AEWs), Kelvin, MRG and IG1 at neighbouring grid points (e.g., over the Guinea coast, over the coast of Senegal and over the north-eastern Sahel). Whilst ER and MJO are unlikely to have significant impacts on daily precipitation, smaller-scale faster waves, such as TDs (not AEW), Kelvin, and IG1, can be generated within the envelopes of these systems (Kiladis et al., 2009) and may be more active and hence can potentially elucidate the aforementioned pattern. As explained before, these faster waves, of course, can significantly influence daily rainfall; 5) While MRG wave activity is more concentrated over the oceanic regions off the coast of Guinea and accounts for approximately 15% of the total grid points, EIG waves are much less common, comprising approximately 9% of grid points, mostly concentrated over the Sahel; 6) IG1 waves make up approximately 10% of all grid points, predominantly concentrated in the northern Sahel (as downstream), namely northern Mali and central-northern Niger and near the coastal regions of Cameroon and Gabon (as upstream). While the former regions experience minimal rainfall, with approximately 50 mm per month in the summer season, the latter regions experience rainfall amounts ranging from over 150 mm per month (near the southern coast of Nigeria) to less than 50 over the coast of Gabon.

Figs. 7.2– 7.4 provide an alternative perspective on the significance of predictors by showing the ranking (lower values are better) distribution across all grid points in the analysis domain. While Figure 7.1 shows only the top-ranked predictor at each grid point, these figures indicate the importance of other predictors. Figure 7.2g consolidates that the TD predictor at the downstream grid point is plausibly the most significant predictor utilised in our study, as it is ranked very high (higher than 4) throughout nearly the entire analysis domain. The low rankings observed over the arid equatorial

Atlantic and the far northeastern Sahel are likely indicative of the weak coupling between the TDs and convection, as TWs that are not convectively coupled cannot be filtered out from the precipitation data. In other words, the high ranks of TDs over the other regions potentially indicate the coupling of AEWs and convection. Meanwhile, the TD-based predictor often yields rankings below 10 at the target and upstream grid points (Figs. 7.3g and 7.4g, respectively) over almost the entire analysis region. As mentioned earlier, it must be noted that the TDs observed over the very southern grid points (south of 5°N) are unlikely to be representative of AEWs.

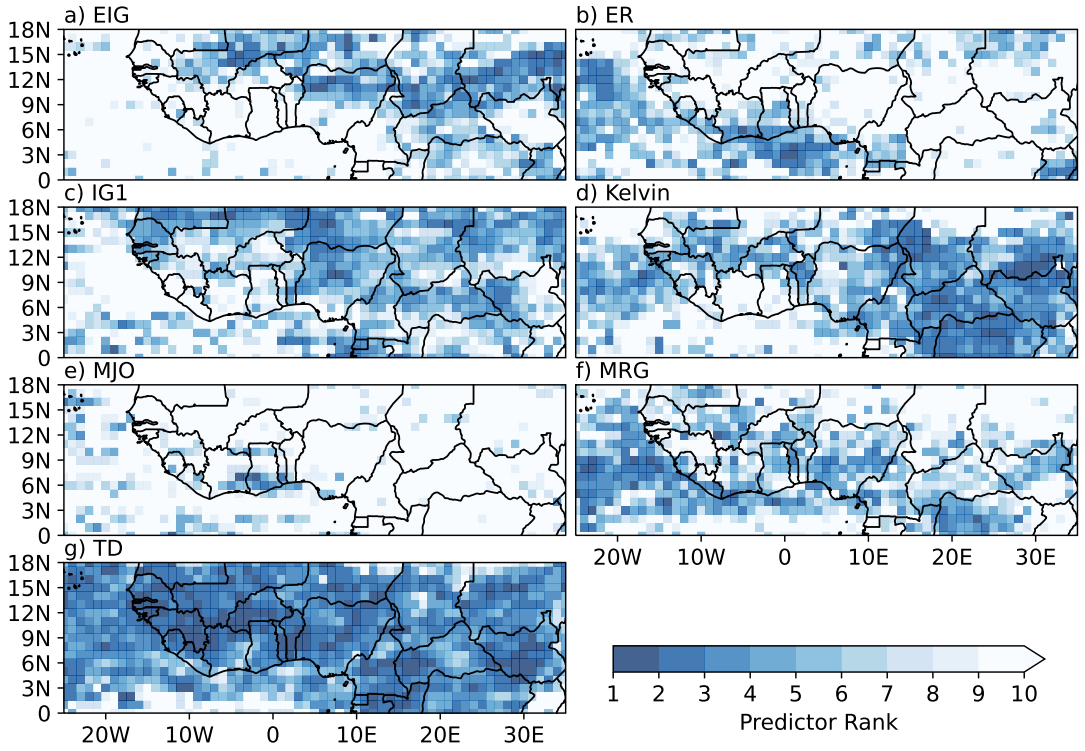



**Figure 7.2:** Rankings of downstream grid point-based predictors based on tropical waves (lower is better) for a) EIG, b) ER, c) IG1, d) Kelvin, e) MJO, f) MRG and g) TD respectively, obtained from the filtered daily rainfall anomaly, are shown for all target grid points during the summer seasons (JAS) from 2007 to 2016 across the analysis domain.

While the Kelvin wave-based predictor at the downstream grid point (Fig. 7.2d) is not as dominant as TD, it still plays a critical role, especially over central Africa and the eastern Sahel. It achieves high rankings (sometimes reaching 1 in certain grid points) across these regions and slightly lower ranks ($\sim 2-5$) over the oceanic area near the coast of Guinea. Kelvin waves couple with convection, which causes them to sometimes follow convective systems to off-equatorial regions. Such patterns may be the possible reasons for the Kelvin wave activity observed in the Sahel in Fig. 7.2d. Unlike the TDs, the Kelvin wave predictors over the downstream grid points seem to show some regional preference; mainly, the grid points over central Africa and the Sahel show high ranks. The absence of identifiable patterns displayed by Kelvin waves at high latitudes (north

of $15°N$) suggests that they are most likely noise, providing additional support for our previous assertion.

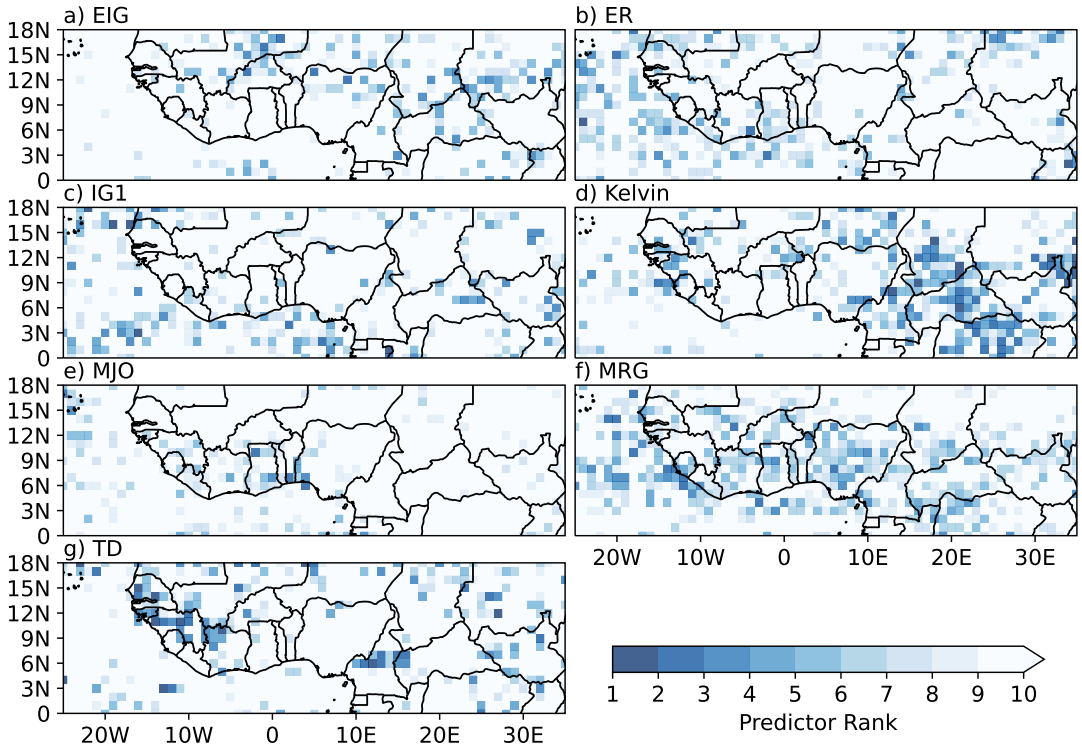



**Figure 7.3:** Same as Fig. 7.2, but for predictors at the target grid point.

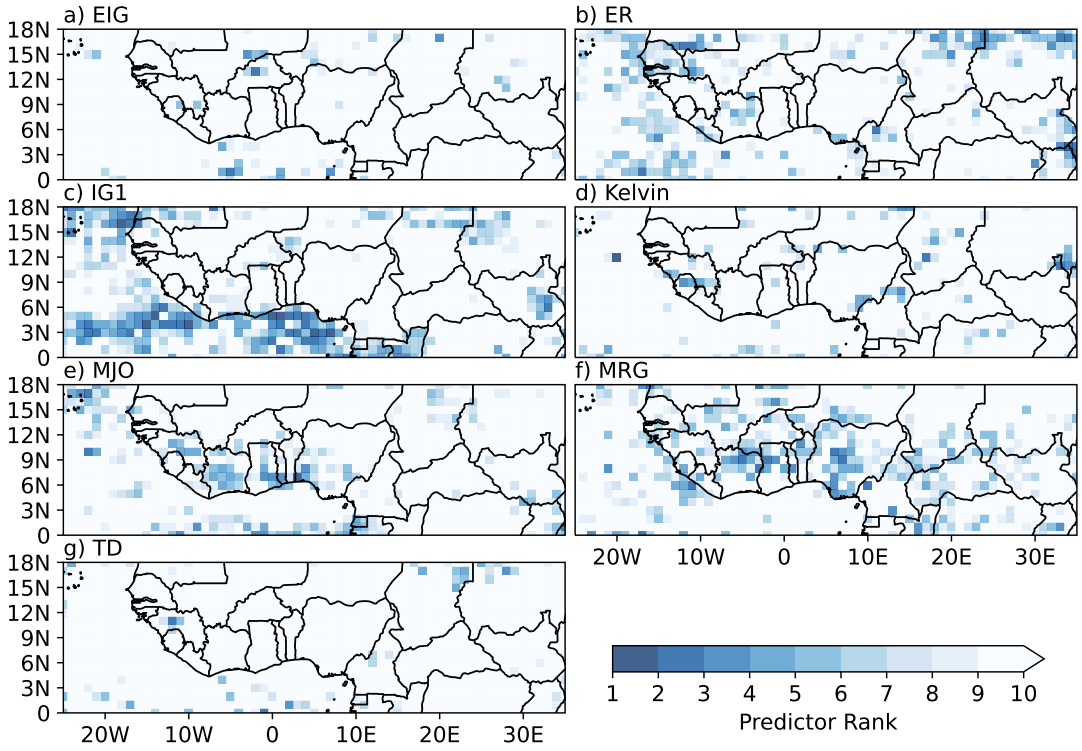



**Figure 7.4:** Same as Fig. 7.2, but for predictors at the upstream grid point.

The IG1 and MRG waves derived from the downstream grid points (depicted in Figs. 7.2c and f respectively) exhibit high ranks throughout the analysis domain. Specifically, the MRG waves are predominant over the central African region, as well as the Sahel and a region over the oceanic region that extends northeastward from the equatorial Atlantic to the Guinea coastline. The ranking of MRG waves observed in the central African region ranges from $2-6$, suggesting that although significant, they are not as crucial as TDs and Kelvin waves. But, as already seen in Fig. 7.1, MRG waves are as active as the TDs and the Kelvin waves over the oceanic region off the coast of Guinea. On the other hand, IG1 waves observed downstream are mainly concentrated over grid points in the land, with higher ranks more common over the Sahel and lower ranks predominant over central Africa (except over Gabon, showing high ranks). Interestingly, the IG1-based predictors at the upstream grid points (Figure 7.4c) exhibit remarkably high ranks, frequently ranking as the top predictor, as shown in Figure 7.1 across the Guinea coast, which aligns with our prior observations. We speculate that the downstream grid points are prominent predictors when information on precipitation at these grid points is vital in forecasting negative impacts, i.e. potential dry conditions over the target grid points. In a recent study, Jung (Jung and Knippertz (2023)) carried out model studies and filtered out IG1s using the WK99 method (Wheeler and Kiladis, 1999), which uses FFTs, and the 2D projection method described in detail in (Yang et al., 2003) on precipitation and dynamical fields respectively. Although the former method effectively identified IG1 signals on the precipitation field, no such signals were detected using the latter method on the dynamical fields. This suggests a lack of association between the two entities, possibly due to the two methods not identifying the same signal. The authors claim that the FFT-filtered IG1s detected in the Sahel are not true tropical waves but rather squall lines associated with mesoscale convective systems (MCS) since these signals are easily picked up by the FFTs on the precipitation field. Our observations are consistent with this conclusion, as we have also detected IG1s in the northern Sahel. However, our results suggest that in addition to the Sahel, such behaviour can potentially be observed over regions surrounding Gabon and the Guinea coast.

While EIG waves are rarely the leading predictor over the study area, according to Fig. 7.1, EIG wave-derived predictors from the downstream grid point are frequently highly ranked and offer vital insights into the precipitation impacting the central to eastern Sahel. This is especially indicated by the high ranks over the largely continuous patch grid points shown over the grid points from Central Mali to Southern Chad, passing through Northern Nigeria and Southern Chad, where MCSs in the form of squall lines dominate rainfall.

The slow-moving wave modes, as anticipated, display mainly low ranks throughout the analysis domain, with the exception of notable high ranks pertaining to ER wave activity (reaching 1) across the coasts and nearby oceanic regions (as shown in Fig. 7.2b at the downstream grid point). We hypothesise that this occurs when low-wavenumber ER waves persist over a region for a sufficient duration, which results in the precipitation

information carried by them from the downstream grid points becoming critical for the forecast in the target grid point.

It is worth considering that the ranks carried by the predictors over the target grid points (Fig. 7.3) are relatively lower compared to those over the downstream grid points. Nevertheless, the observed patterns are somewhat comparable to those in the downstream predictors for some wave modes, like the MRG and Kelvin waves.

These results show that patterns formed only by the top predictors are likely to explain, at most, the first-order influences of TWs in the daily rainfall modulation. In the predictor selection procedure, we do not impose strict limits on the number of predictors to be selected, except for the threshold, which determines the selection of those whose coefficients exceed it. Despite this, the predictor selection procedure chooses approximately 13 to 16 predictors out of the available 21 to train statistical models at all grid points (not shown). Restricting the selection to only the highest predictors when training models markedly decrease their skill (not shown). This suggests that the relationship between different TWs and, in turn, their combined interaction with daily rainfall accumulation is far from linear.

## 7.2 Modulation of daily rainfall by tropical waves

From the findings in the previous Section, specific regions where the wave activity is concentrated and relevant to practical predictability are identified (refer to Figs. 7.1 and 7.2-7.4). The wave activities computed over these regions are then utilized for the computation of phases and amplitudes (for more information, see section 5.2.3). Next, we identify the dates during which the wave activities defined in these areas were in each of the eight phases in order to compute composites. Note that an amplitude threshold has not been set to identify phases, as it has not produced significant differences between tests conducted with thresholds set at 0.0, 0.5 and 1.0 standard deviations. Furthermore, this approach enables the inclusion of more data points, resulting in a more robust result.

In the following subsections, we attempt to elaborate on rainfall modulation by tropical waves using rainfall anomaly composites, which are based on the phases of each wave mode. We opt to discuss each wave mode in detail due to the varied regions where the phases of these waves are determined (indicated by dashed regions in Figs. 7.5-7.11), with the aim of maximizing their respective wave activity.

### 7.2.1 Tropical Depression (TD)

The composite of rainfall anomaly during phases of tropical depressions (TDs) over summers spanning from 2007 to 2016 is presented in Figure 7.5. The phases are defined according to computed wave activity inside the dash-boxed area. Although TDs are active throughout most of the analysis region, as demonstrated in Fig. 7.2g, and Fig. 7.1g highlights significant clusters of TDs in the Eastern Sahel and central Africa, we chose to

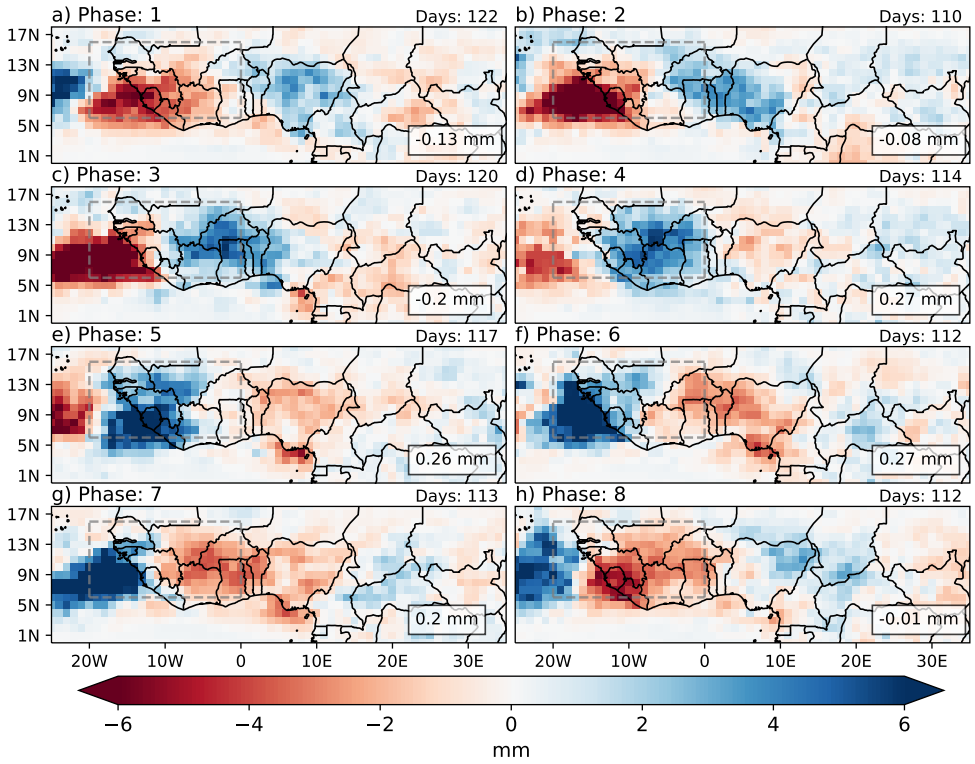



**Figure 7.5:** Composite of the daily rainfall anomaly during the Phases of tropical depression (TD) for the summer seasons (JAS) spanning from 2007 to 2016 over the analysis area (Note that this differs from the composite of TD-filtered precipitation anomaly field). The Phases are defined in the region indicated by the dashed box. The total number of days in the composite is shown in the upper right of each panel. The dashed region in each panel indicates the region of high TD-wave activity. The area-averaged rainfall anomaly inside this box is shown in the bottom right corner. Blue (red) shading indicates a positive (negative) anomaly in rainfall.

focus on the region depicted in Fig. 7.5 for two main reasons: Firstly, our interest lies in TDs acting as proxies for AEWs, and thus, signals south of 5° N lack meaningful value. By limiting our scope to a specific area, we can more effectively examine and analyse the impact of AEWs in this region. Secondly, there is a higher climatological activity of African Easterly Waves (AEWs) in the Western Sahel compared to the Eastern Sahel due to the stronger African Easterly Jet (AEJ) prevailing in the West during the summer season (Fink and Reiner, 2003).

Fig. 7.5 clearly shows that the rainfall anomalies vary significantly depending on phases of TD within the dashed region. As expected, the region is predominantly dry during Phase 1, with positive anomalies to the West and the east. Positive rainfall anomalies can be seen propagating inward from the east in Phase 2, indicating westward propagating AEWs. By Phase 3, which is a transition phase, the positive and negative anomalies occupy the region roughly in halves. However, the negative anomalies are of higher magnitudes than the positive ones. This is likely indicative of the strong coupling between TDs and rainfall. As we move into Phase 4, the region experiences an increase in precipitation. This is supported by the positive area-averaged precipitation amount

(refer to Figure 7.5d, bottom-right). In Phase 5, the intensity of the positive anomalies continues to increase, with a concentration over the western coast of West Africa and maximum values exceeding 6 mm per day. Simultaneously, the eastern edge of the boxed region begins to show a decrease in precipitation levels, with negative anomalies approaching from the east. By Phase 6, the positive anomalies have moved further west, and more eastern grid boxes display negative anomalies. Phase 7 marks another transition phase but with an opposite orientation compared to Phase 3, and the region is shared almost equally by positive and negative anomalies. Phase 8 reveals the positive anomaly exiting through the West, resulting in a predominantly dry region, which is also evidenced by the slightly negative area-averaged rainfall amount.

### 7.2.2 Kelvin waves

Fig. 7.6 displays the composite of rainfall anomalies during the phases of the Kelvin wave. The area marked with dashed lines, covering several parts of the eastern Sahel and central Africa, visible in all panels, is the region selected for computing the phases and amplitudes. This particular area is chosen on the basis of Figs. 7.1 and 7.2d to maximize Kelvin wave activity detected. Additionally, we limit ourselves to latitudes below 15° N since, theoretically, Kelvin waves are equatorially confined.

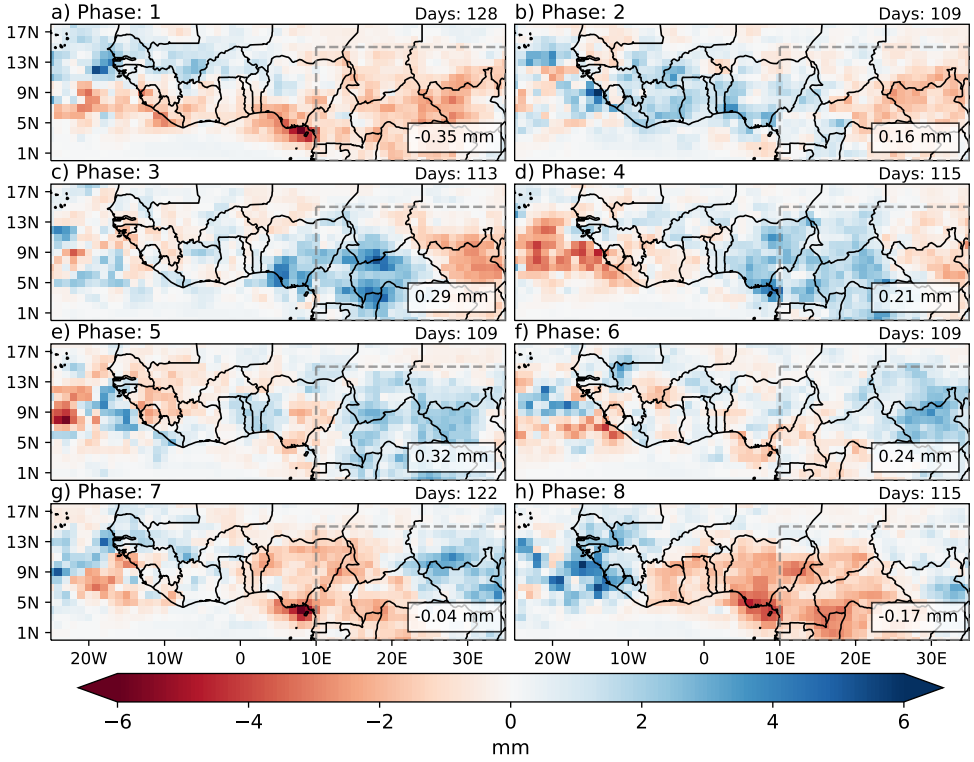



**Figure 7.6:** Same as Fig. 7.5 but for Kelvin wave.

During Phase 1 (Fig. 7.6a) of the Kelvin waves, the area within the dashed box is predominantly arid, with a notable minimum of around −2 mm in multiple grid boxes across South Sudan. This is further supported by the negative area-averaged

rainfall (−0.35 mm) displayed in the bottom-right corner. Slightly positive precipitation anomalies are observed in the region's west in Phase 2, indicating eastward propagation of Kelvin waves. These are also seen extending further west beyond its borders. Nonetheless, the vast majority of the region remains dry. During Phase 3 (transition period), the rainfall anomalies increase in strength, reaching up to 5 mm over certain grid points. These anomalies spread further inward across the whole western half of the domain while the eastern half remains dry. Phase 4 exhibits patterns akin to Phase 3, with the exception of greater inland penetration of positive anomalies, particularly over the northern DRC. Furthermore, both positive and negative anomalies appear to be more diffused.

By Phase 5, the entire region enclosed by the dashed box displays positive rainfall anomalies, further confirmed by a positive area-averaged rainfall of (0.32 mm). During Phase 6, negative anomalies propagate inwardly from the west. However, several regions inside the dashed box exhibit positive anomalies, exhibiting higher values ($\sim 2$ mm), particularly towards the east in South Sudan. Phase 7, being a transition Phase, shows an even mix of positive and negative rainfall anomalies but opposite to what was observed in Phase 3. The negative anomalies observed during this Phase are noted to expand towards the west, reaching as far as Ghana, after which the patterns become mixed. The negative anomalies continue to propagate inwards and intensify in Phase 8, with values around $\sim -3$ mm, along the northern border of the Congo. Nonetheless, the far eastern edge of the region still shows positive anomalies.

It seems that the Kelvin wave triggers significant convective activity upon reaching the western edge of the dashed area in Phase 3, leading to suppression of rainfall towards the east and creating an east-west dipole. This pattern is also evident in Phase 4. Upon the Kelvin wave's exit from the marked region in Phase 7, the opposite occurs. The dipole structure results in a much shorter wavelength of the wave in phases 3, 4, 6, and 7 compared to the others. It is worth noting that during Phases 1, 7, and 8, rainfall anomalies as low as approximately -5 mm are observed along the western coast of Cameroon. However, these anomalies lie outside of the selected region. Positive anomalies are observed during Phase 4, exhibiting similar patterns. The reason for such patterns may be multi-folded; this region receives large amounts of precipitation, largely orographically induced due to topography around $\sim 1800$ m (see Fig. 5.2). Tropical waves, filtered from precipitation data utilizing FFT-based techniques (similar to those employed in this study), are prone to display robust signals across regions characterized by high precipitation. This outcome transpires due to the inevitable overlapping between the wavenumber-frequency realms of such signals and tropical waves. Nevertheless, tropical waves could also couple with these patterns, potentially instigating or hindering precipitation. However, the precise reason remains unclear at this point and requires further investigation in the future.

### 7.2.3 Mixed Rossby-gravity (MRG) waves

Fig. 7.7 displays the rainfall anomaly composite during MRG phases in the summer of 2007-2016. The dashed area used for computing the phases and amplitudes of MRG waves is similar to the one used for TDs. We chose this region because of the high MRG activities in Figs. 7.1 and 7.2f. As MRG waves and TDs share some characteristics like frequency and direction of propagation, overlaps of a certain degree may be observed in the rainfall anomaly patterns during some phases of MRG waves.

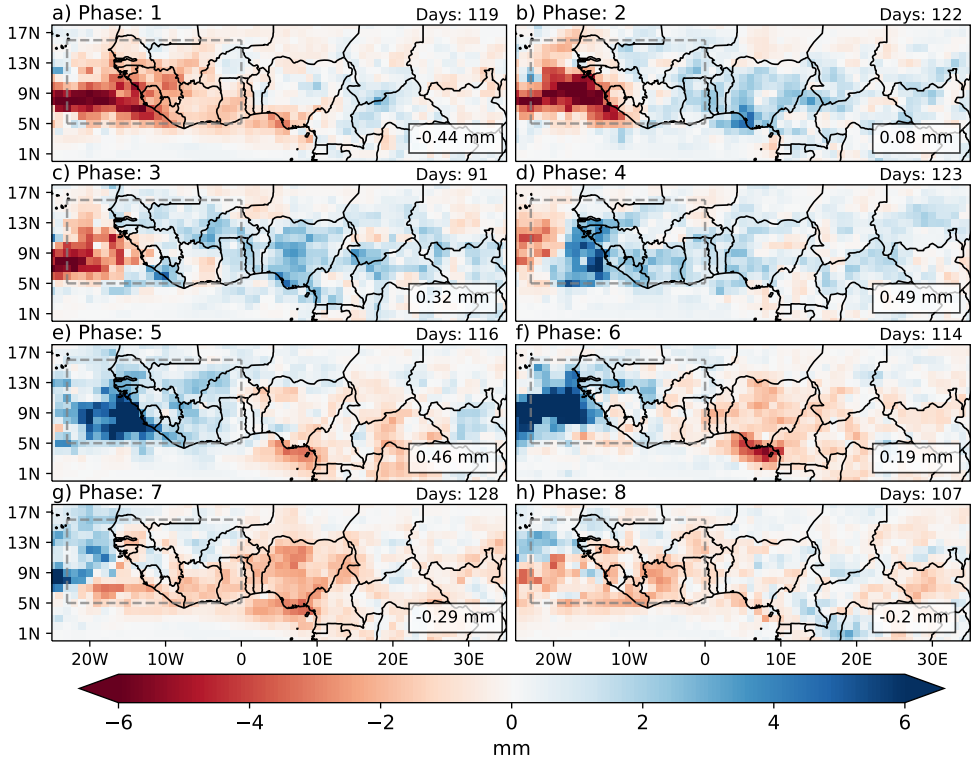



**Figure 7.7:** Same as Fig. 7.5 but for MRG wave.

The region within the dashed box comprises mainly grid points displaying negative rainfall anomalies in Phase 1, reaching values as extreme as −6 mm per day over the oceanic area off the coast of Guinea and nearly zero over the land regions in Côte d'Ivoire. The negative area-averaged precipitation (−0.44 mm) indicated in the bottom-right corner further confirms the 'dryness' of Phase 1. The eastern portion of the area exhibits slightly positive precipitation anomalies during Phase 2. In contrast, strongly negative precipitation anomalies characterize the western side with a more considerable north-south extent than Phase 1. The excursion of positive anomalies through the eastern edge of the region indicates the westward propagation of the MRG waves. While the general patterns of Phases 1 and 2 are similar to those of the TDs, the amplitudes of the MRG anomalies (especially the positive ones) are generally smaller than those of the TDs. Furthermore, MRG wave anomalies demonstrate a more spatially diffused structure than TDs. Phase 3 has more grid points showing positive rainfall anomalies than Phase 2, but the western edge is still largely negative. Phase 3 of MRGs differs

notably from that of TDs with regard to the gradient between positive and negative anomalies.

In contrast to TDs, where grid points indicating negative and positive rainfall anomalies were sharply separated into the west and east, respectively, the separation in Phase 3 of MRGs is smoother. In Phase 4, the marked region mostly shows positive rainfall anomalies at most grid points, with a few negative anomalies in the west. Positive rainfall anomalies reach a maximum of 6 mm per day but are limited to roughly $3-4$ grid points along the coast of Sierra Leone. The precipitation structure also appears less organised than TD in Phase 4. Nearly all grid points within Phase 5 depict positive rainfall anomalies. The magnitude of precipitation over land areas is generally low, but it exceeds 6 mm per day along the coasts of Senegal, Sierra Leone and Gambia. The southwest-northeast direction alignment of rainfall structure suggests a strong coupling between precipitation and MRG waves in this region during the summer season, as evidenced by the sizeable amounts of rainfall also aligned in the same fashion(refer to Fig. 6.1's top panel's darkred isohyets). In Phase 6, the southwest-northeast structure strengthens as the positive anomalies continue to move towards the west. Simultaneously, the eastern side of the region begins to display minor negative anomalies. High levels of rainfall are also detected along the Cameroon coast, exhibiting parallels to both TDs and Kelvin waves. However, while rainfall patterns within Phase 7 (transition) took on contrasting patterns to those viewed in Phase 3 for the former two, MRG waves illustrate a rainfall gradient oriented along the northwest-southeast axis matching the general rainfall pattern in this region. By Phase 8, negative rainfall anomalies are observed in most grid points within the region, with the exception of some positive values on the northern side. Nevertheless, no discernible pattern other than this is observed in Phase 8.

It is interesting to note that the signals are largely confined to the oceanic regions like in Fig. 7.2f, indicating that the convergence (in the case of positive anomalies) and divergence (in the case of negative anomalies) patterns of MRG waves are more relevant over the oceanic region compared to the land regions. It is noteworthy that the structures of rainfall anomalies exhibited during the MRG phases bear a resemblance to the phases of the AEW-MRG 'hybrid' wave outlined by Cheng et al. (2019), suggesting a potential coupling between the two waves.

### 7.2.4 Madden-Julian Oscillation (MJO)

The composite of rainfall anomalies during the phases of MJO in the summer seasons between 2007 and 2016 is illustrated in Fig. 7.8. Due to the limited predictive power of MJO-based predictors at only a few grid points (refer to Figs. 7.1, 7.2e and 7.4e) lacking any discernible region-specific pattern, we have chosen to perform phase and amplitude calculations for MJO across a large region encompassing West Africa and the nearby oceanic area as indicated by the dashed box region in Fig. 7.8. As it is highly improbable for MJO phases to impact daily rainfall patterns significantly, none

of the phases exhibit distinguishable rainfall patterns as observed in previous cases. Nevertheless, some discernible differences between the wet and the dry phases can be observed.

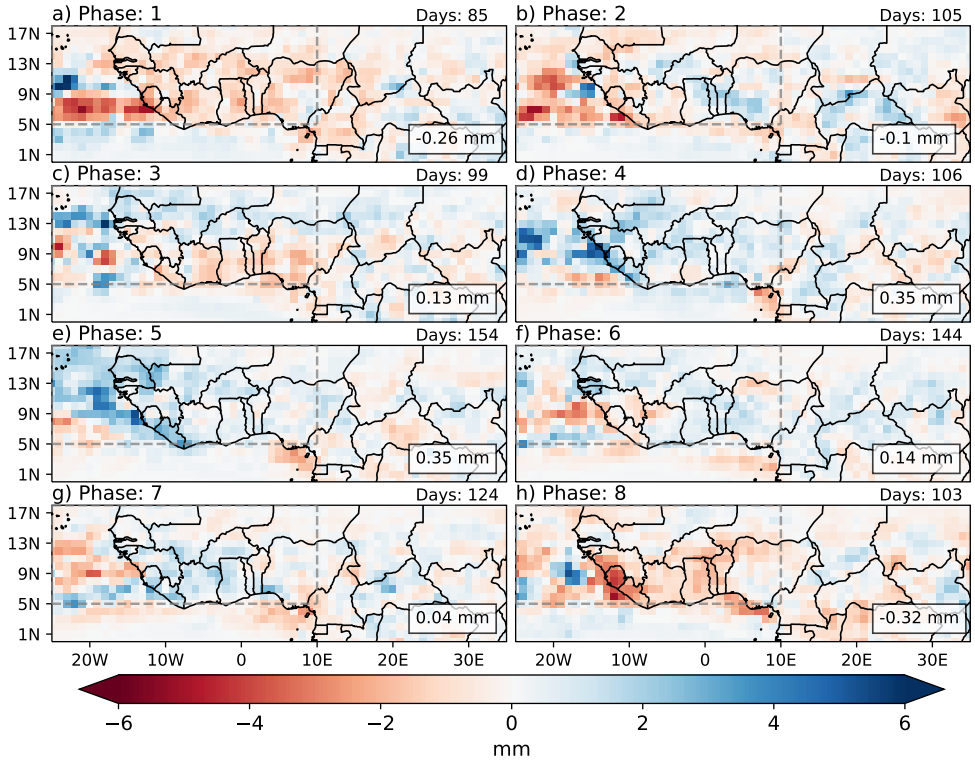



**Figure 7.8:** Same as Fig. 7.5 but for MJO.

The land-based grid boxes encompassed in the dashed box mainly show negative rainfall anomalies in Phase 1 of MJO but are largely diffused. The coast of Sierra Leone and parts of the oceanic region show anomalies as low as $\sim -5$ mm. Some positive rainfall anomalies are also observed in the far West over the ocean. Mild positive rainfall anomalies are detected in certain grid boxes located in the Central-Eastern part of the dashed region, inclusive of southern Nigeria, Benin, Togo, northeastern Ghana, Niger and southern Mali. Conversely, other land-based grid boxes in the region display minor negative anomalies. With the exception of a few grid points along the western coast, the oceanic region is mainly arid, with negative anomalies up to -5 mm daily. The dryness of Phases 1 and 2 is also depicted by the negative rainfall average displayed on the lower right section of Figs. 7.8a and b. Phase 3 of the MJO exhibits a separation of positive and negative anomalies in rainfall from north to south over land. Specifically, the northern half (above $\sim 11^\circ$ N) depicts positive anomalies, while the southern half demonstrates negative anomalies. However, no discernible pattern of rainfall anomalies can be identified over the oceanic grid points in the region. Phases 4 and 5 exhibit comparable patterns, with almost all grid points in the region presenting positive anomalies. Nevertheless, the grid points situated along the western coast and in the ocean illustrate a more coherent structure of positive anomalies in Phase 5

compared to Phase 4. Furthermore, Phases 4 and 5 show positive area-averaged rainfall amounts (0.35 mm). Phases 6 and 7 demonstrate a gradual movement of negative rainfall anomalies from the western region, suggesting an eastward propagation of MJO. This is corroborated by the decreasing mean rainfall, as depicted in the bottom left corner of Figs. 7.8f and g. In Phase 8, the land region exhibits predominantly negative rainfall anomalies, indicative of a dry Phase, with the grid points in Sierra Leone registering a minimum of approximately $-4$ mm per day.

### 7.2.5 Westward inertio-gravity (IG1) waves

Fig. 7.9 presents the rainfall anomaly composite during the Phases of IG1. The phases and amplitudes are computed in the dashed region covering parts of the western to central Sahel. This location was chosen as it has been identified through the analysis of Figures 7.1 and 7.2c presented in the previous chapter that there is a significant presence of IG1 wave activity.

While a few grid points in the eastern area of the region (northern Nigeria and central Niger) exhibit marginally positive rainfall anomalies ($\sim -0.5$ mm per day), the majority of grid points in the area infer negative anomalies during the IG1 wave's initial Phase. This is further confirmed by the negative average daily precipitation ($-0.04$ mm) indicated in the lower left-hand corner of Fig. 7.9a. An excursion of notably positive anomalies can be observed in the eastern sector of the demarcated area during Phase 2. However, in comparison to the other wave modes, such as the TDs and MRG waves, IG1 exhibits much less significance in daily precipitation modulation, as indicated by the low magnitudes of rainfall anomalies. Phase 3 exhibits a comparable trend with Phase 1, wherein nearly the entire region encountered arid conditions. However, during Phase 4 (the onset of the wet Phase), positive rainfall anomalies emerge in the central region of the indicated area, with the western portion remaining dry. At the peak of the wet Phase (Phase 5), the positive anomalies that are seen in Phase 4 have shifted westwards, suggesting the westward propagation of the IG1 waves.

While the northern half of the marked area exhibits almost zero rainfall anomalies, which is anticipated due to its climatological aridity, the southern half displays positive anomalies of up to 2 mm per day. Phase 6 demonstrates the continued westward propagation of these positive anomalies, with the eastern edge experiencing drier conditions. Phase 7 displays characteristics of a transition season, with the western portion of the marked area exhibiting rainfall anomalies as high as 4 mm per day and negative anomalies evident in the east. However, the northern region in both areas remains weak, with values close to zero. Phase 8, marking the beginning of the dry Phase, displays a slight negative anomaly in the central region, while the Western area continues to remain marginally wet.

The lack of wave-like activity during the IG1 phases may indicate less coupling with convection compared to other wave modes such as TDs, MRG and Kelvin waves. It

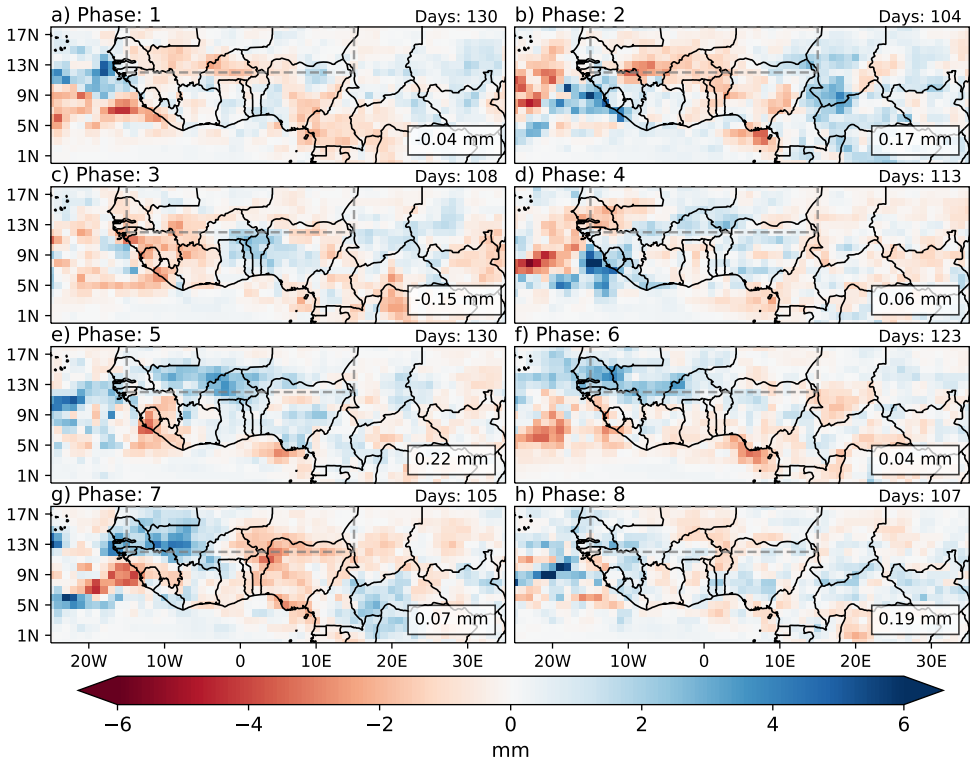



**Figure 7.9:** Same as Fig. 7.5 but for IG1 wave.

is also possible that the IG1 signals identified are rapid-moving MCSs (as evidenced by the work of Jung and Knippertz (2023)). Identifying these at a daily resolution can be challenging, which could account for weak but discernible patterns observed during select phases (such as the clear westward propagation of positive rainfall anomalies seen from Phases 4 to 6).

### 7.2.6 Equatorial Rossby (ER) waves

The composite of rainfall anomalies during the ER wave's phases for the summer season between 2007-2016 is displayed in Fig. 7.10. Figs. 7.1 and 7.2b are used to determine the locations of significant ER-based predictors in rainfall forecasting. We then calculate the Phases and amplitudes of the ER wave inside the region indicated by dashed lines in Fig. 7.10.

Most of the region identified by the dashed lines displays negative rainfall anomalies, with some as low as $-4$ mm during Phase 1, except for a few small patches of slightly positive anomalies located along the western coast of Guinea. This small wet patch surrounded by drier conditions could be due to orographically triggered precipitation in the Guinea highlands. The average daily rainfall for this Phase is $-0.25$ mm. Positive rainfall anomalies are visible over parts of Southern Mali, Burkina Faso, and Southern Niger in Phase 2, while the rest of the marked area remains mostly dry. Large areas of positive rainfall anomalies noted across central Africa during Phase 2 arrive at the designated area by Phase 3, demonstrating the westward propagation of the ER waves.

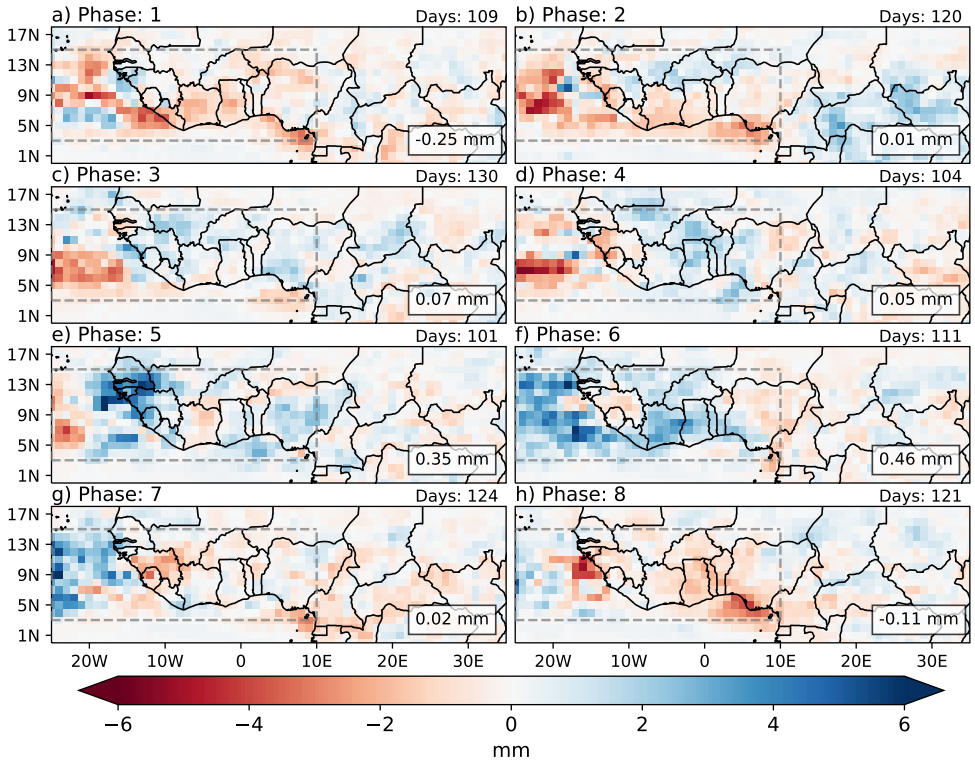



**Figure 7.10:** Same as Fig. 7.5 but for ER wave.

The rainfall anomalies spotted over Nigeria, as well as most of the regions referenced earlier, can be attributed to this. The moist conditions shift marginally towards the West and are now centred over the middle of the demarcated region in Phase 4, while the areas in the vicinity of the Guinea highlands continue to experience dry conditions. Rainfall anomalies up to 6 mm may be perceived across the Guinea highlands and the nearby areas at the peak of the wet Phase (Phase 5). Except for small pockets of drier conditions over Côte d'Ivoire and parts of southern Mali and northern Nigeria, all areas show positive rainfall anomalies. Nevertheless, the oceanic region situated in the westernmost part of the delineated region continues to remain dry. Phase 6 indicates the greatest amount of area-averaged rain (0.46 mm per day), which can be credited to the significantly positive rainfall anomalies observed in large areas of the marked region. Nevertheless, an area of drier anomalies is visible over Guinea, most likely caused by the elevated altitudes of the Guinea highlands. The anomalous observations detected over Guinea during the wet Phases 4, 5, and 6, relative to nearby areas, may suggest that the impact of the ER wave is less significant than orography-related features in inducing or suppressing rainfall in this region. Finally, Phases 7 and 8 demonstrate further westward propagation of positive rainfall anomalies and the approach of drier conditions from the East of the demarcated region.

### 7.2.7 Eastward inertio-gravity (EIG) waves

Based on Figures 7.1 and 7.2a, it appears that EIG waves are most active in the central to eastern region of the Sahel and potentially contribute, at least in part, to squall line-dominated rainfall events. To investigate the activity of EIG waves over these regions during different phases of the wave, we performed the phase-amplitude calculation in the outlined region shown in Fig. 7.11.

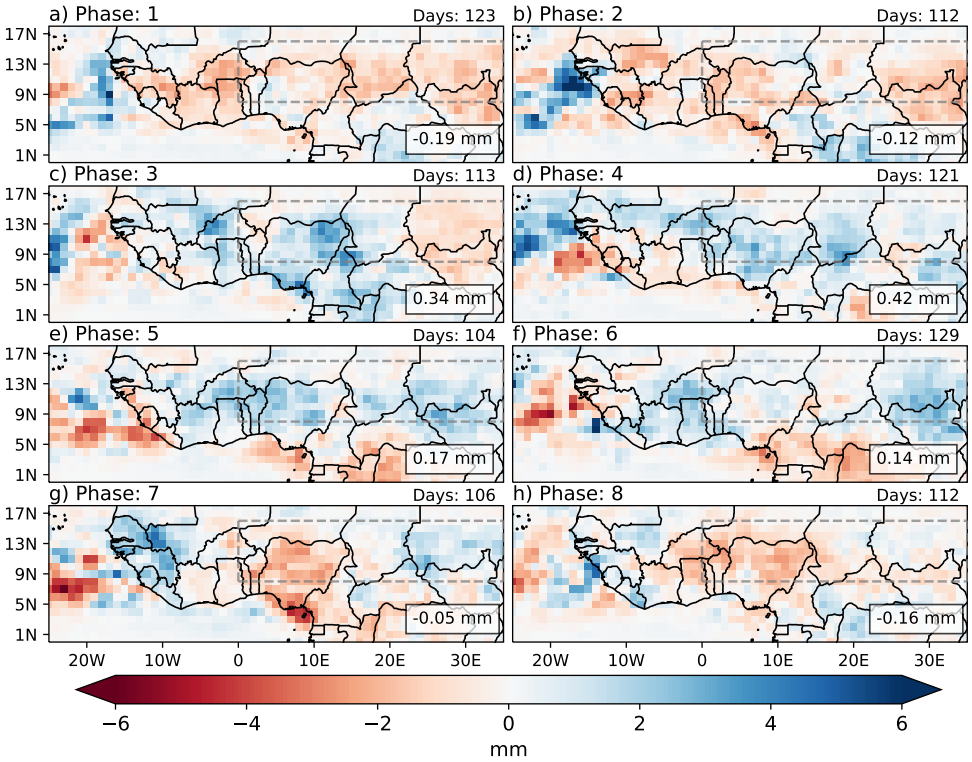



**Figure 7.11:** Same as Fig. 7.5 but for EIG wave.

Negative rainfall anomalies are observed over nearly all grid points in the region throughout Phase 1. However, the slightly positive anomalies observed over northern Benin and parts of western Nigeria are an exception. As would be expected, Phase 1 is also the driest Phase, with an average rainfall of −0.19 mm. Except for some regions in southern Niger and southern Chad, the entire demarcated area experiences dry conditions in Phase 2. Although the eastern section of the region remains arid, positive precipitation anomalies can be observed over the central to western sections. The Central African region, located south of the marked area, also exhibits wetter conditions during Phase 3 of the EIG wave. By Phase 4, however, the entire demarcated area displays wet conditions, extending westward to the Atlantic Ocean. Consequently, Phase 4 shows the highest average rainfall (0.42 mm per day). The rainfall anomaly field patterns remain broadly the same in Phase 5 but appear to be less intense than in Phase 4. The positive rainfall features in Phases 4-6 are concentrated over the Sahel region, perhaps partly triggering the MCS-dominated rainfall. By Phase 6, negative rainfall anomalies are developed over southern Chad and parts of southern Niger, with the east section of the

marked region remaining moist. Entering the region from the west in Phase 7, negative rainfall anomalies as low as −3 mm per day suggest the eastward propagation of the EIG waves, while the eastern edge of the marked region continues to remain wet. The patterns remain almost the same in Phase 8 but with a lower intensity. The absence of a distinct wave-like pattern and lower levels of rainfall anomalies in comparison to more powerful waves, like the TDs and Kelvin waves in the Sahel region, suggests weak convective coupling of EIG waves.

## 7.3 Modulation of forecast skill by tropical waves

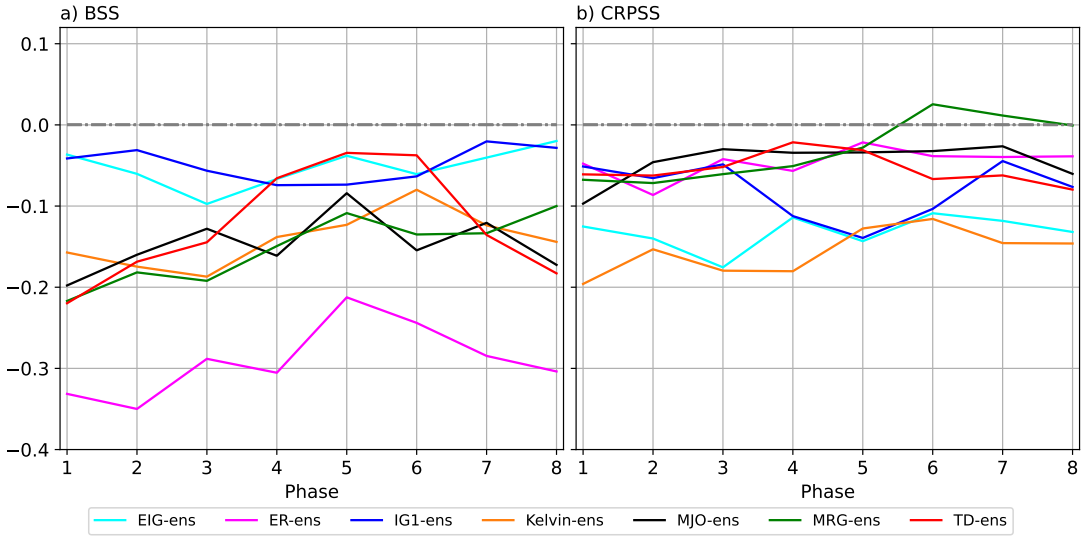



**Figure 7.12:** The average Brier skill score (BSS) and Continuous Ranked Probability Skill Score (CRPSS) for the summer seasons of JAS between 2007-2016 are displayed for the ENS-raw ensemble forecast compared to the EPC15-ensemble forecast (reference) during the different Phases of EIG (cyan), ER (pink), IG1 (blue), Kelvin (orange), MRG (green) waves, MJO (black), and TDs (red).

Now that it has been confirmed that the rainfall anomalies are influenced by the phases of the tropical waves (some waves clearly more than others), we proceed to analyse how the skill of NWP model-based forecast is modulated, namely the ENS ensemble forecast. In addition to assessing the skills of these forecasts in predicting the daily amount of rainfall, we will also analyse the modulation (if any) of this forecast in predicting the occurrence of rainfall at a threshold of 0.2 mm per day.

Fig. 7.12 illustrates the deviation of the mean Brier Skill Score (BSS) and Continuous Ranked Probability Skill Score (CRPSS) of the ENS forecast relative to the EPC15 ensemble forecast (reference) during the phases of tropical waves in the summer seasons of 2007-2016. The phases have been identified, as explained in the previous section. Similar to the analyses in the preceding section, we refrain from setting any amplitude threshold when defining the phases, as it did not reveal any significant differences (see Figs. A4 and A5).

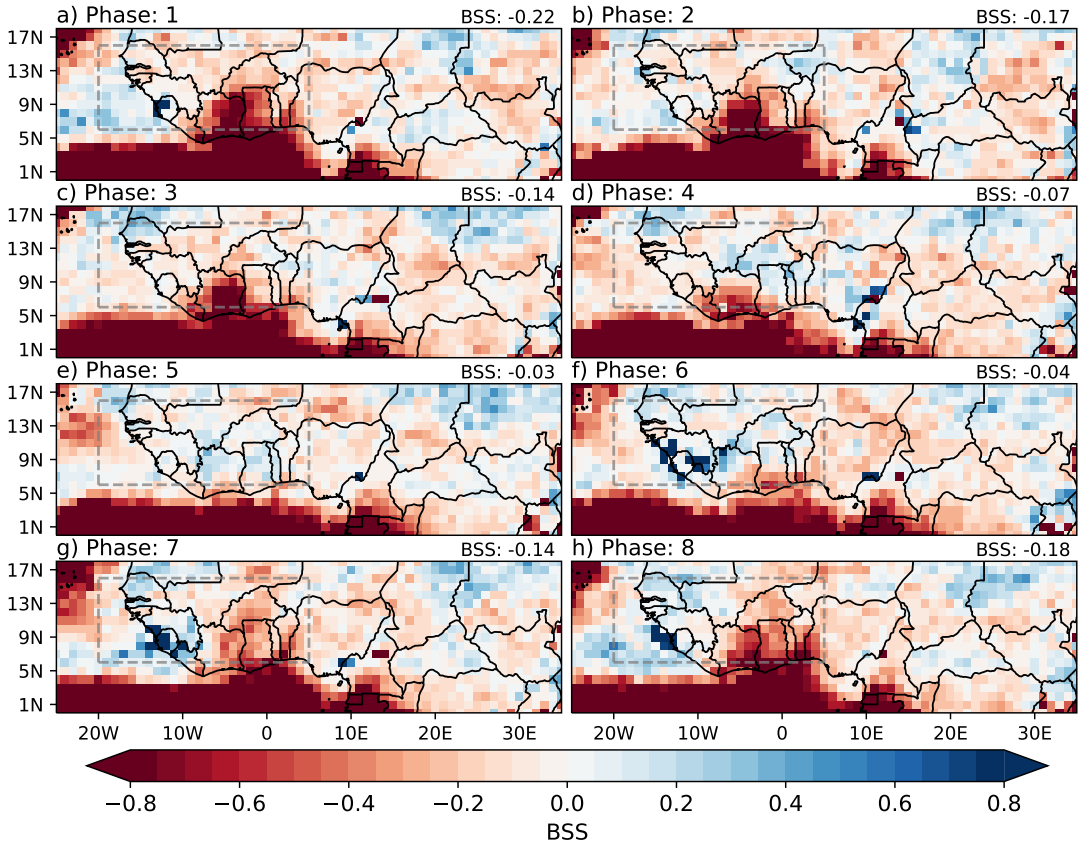



**Figure 7.13:** Modulation of the Brier Skill Score (BSS) of the raw ensemble forecast (ENS) versus the EPC15 ensemble forecast (EPC15-ensemble) over tropical Africa during Phases of the IG1 wave defined within the region marked by dashed lines. The average BSS within the demarcated area (blue solid line in Fig. 7.12a) is displayed in the top right corner of each figure.

(1) All tropical waves, except for IG1 and, to some extent, EIG waves, significantly affect the BSS of ENS forecasts, as illustrated in Fig. 7.12a) (solid lines). The BSS of the ensemble forecast (ENS) displays distinct peaks during the wet Phases (4-6) of TDs, MJO, Kelvin, MRG and ER wave modes. However, a notable decline in the BSS is seen during the wet phases of IG1, and no distinct pattern is observed between the BSS of the ENS forecast and EIG phases. While it may be possible that IG1 waves modulate the rainfall occurrence skill during dry phases, it should be noted that the fluctuations in BSS throughout the various phases of EIG waves are substantially lesser than other wave modes, implying the absence of any apparent modulation by EIG. Some possible explanations for the pattern observed in IG1 are as follows: It can be inferred from Fig. 7.13 that the low BSS in Phases 1, 2, 7 and 8 is primarily due to positive BSS values in grid points over southern Mauritania, northern Senegal and partly northern Niger. According to the dark red isohyets depicted in Figs. 6.1a-c, it is apparent that these areas have very low rainfall during JAS. However, from the investigations of Knippertz et al. (2003), it is clear that while these regions are generally climatologically dry, the occurrence of rare explosive precipitation events cannot be completely ruled out.

During such instances, fast-moving AEWs (which share the same signal as convectively coupled IG1 waves as pointed by Jung and Knippertz (2023)) can, in theory, interact with these convective events, leading to precipitation. In addition, Knippertz (2003) examine precipitation events over northwestern Africa caused by tropical-extratropical interactions (TEIs), which may be better depicted in the ENS ensemble forecasts. It should also be added that, due to the arid nature of the demarcated area, utilizing convectively coupled IG1 waves to characterize phases may not be a reasonable approach compared to using stronger waves like TDs. Such an approach could also contribute to the occurrence of the aforementioned cases during the dry phases of IG1 instead of the wet phases. Some or all of these factors, when combined, may contribute to a better skill in the ENS ensemble forecast, leading to a positive BSS in these areas, resulting in an increased mean BSS in the demarcated region during the dry phases of IG1.

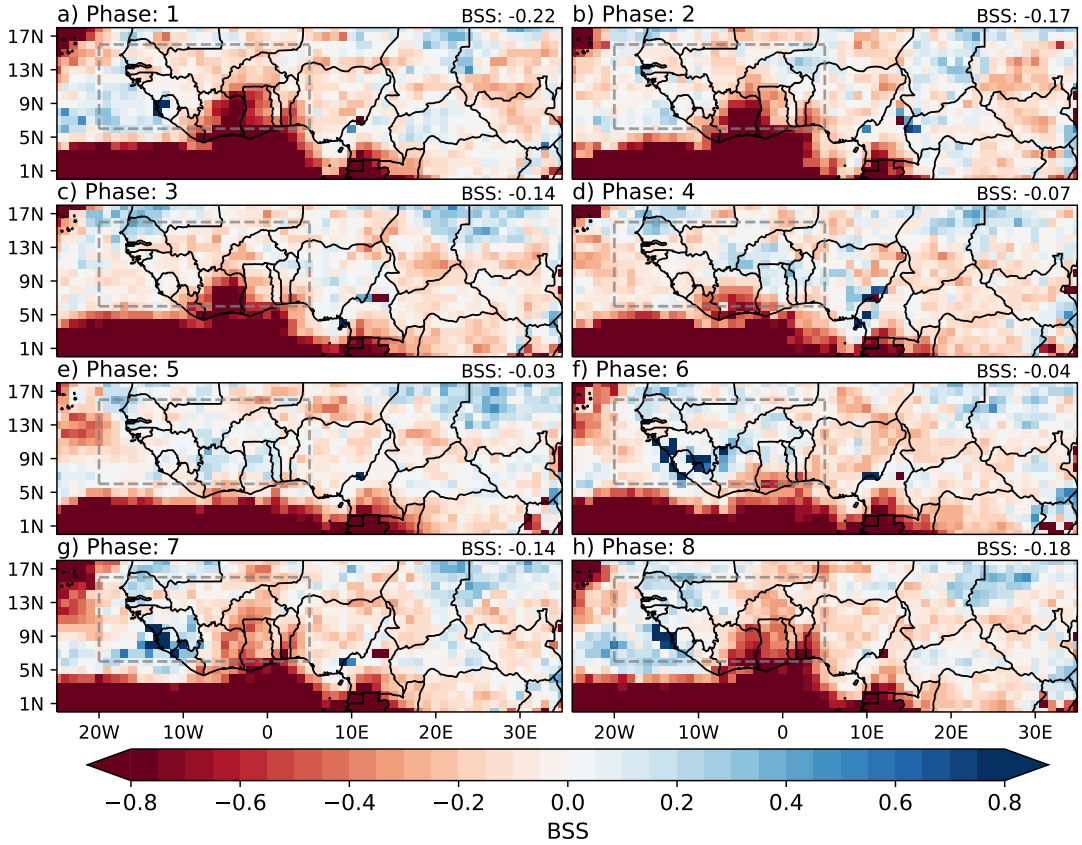



**Figure 7.14:** Same as Fig. 7.13 but for TDs.

(2) Among the tropical waves that influence the BSS of the ENS ensemble forecast, TDs show the most pronounced increase during the wet phases, with a difference of $\sim -0.19$ between the highest (during Phase 5) and the lowest (during Phase 1) BSS. The BSS per grid point across the analysis domain during TD phases is illustrated in Figure 7.14. It can be deduced that the elevated BSS of ENS forecast in the wet phases of TD (Phases 4, 5 and 6) primarily results from the enhanced BSS (see Fig. 7.14e) in Côte d'Ivoire. It is noteworthy that this 'kink' of enhanced predictability higher than

the adjacent areas is observed multiple times and by various means in this region (e.g. Fig. 6.1a-c, Fig. 6.7a-c, ...). This can be explained by the fact that the MCSs initiated over the Cameroon highlands and the Jos Plateau are typically short-lived and dissipate before they reach the border of Ghana-Côte d'Ivoire as detailed in Lafore et al. (2017). Therefore, from a climatological perspective, this region experiences less precipitation and is comparatively drier than the surrounding areas. However, with the presence of a strong driver such as the AEWs (or TDs in general), the MCSs can be active for much longer and potentially propagate much further. As the BSS score is relative to a reference forecast (in this case, the EPC15 forecast), the positive BSS in Phase 5 can be explained as follows: the ENS forecast predicts precipitation in JAS more often (false positive) than the climatology (EPC15) due to modulation by AEWs over this region and generally underperforms. However, the BSS improves during the wet phases of the TDs due to a potential increase in rainfall occurrence in the EPC15 forecast. Similar features can also be observed in the case of MRG, although not as contrasting as in TDs, when the kink over Côte d'Ivoire is significantly less prominent during the wet phases (see Fig. A3).

(3) Considering the wave modulation of the CRPSS (see Fig. 7.12b) in the ENS forecast, it is noticeable that comparable patterns to the BSS can be found, albeit with slight variations. Improved CRPSS is observable during one or more of the wet phases of TDs (Phase 4), Kelvin (Phases 5 and 6), MRG (Phase 6), and ER (Phase 6) waves. Similarly to the BSS, EIG waves do not show any discernible effect on the CRPSS. Like the BSS, there is a significant drop in the CRPSS during the wet Phases $(4-6)$ of the IG1 wave. However, it is noteworthy that the CRPSS demonstrates no modulation with the phases of the MJO, unlike the BSS. The CRPSS reaches its highest value during Phase 6 of the MRG wave (0.03) and its lowest during Phase 1 of the Kelvin wave (-0.2). The variance of the CRPSS between the different phases of the waves is significantly smaller than that of the BSS, suggesting that tropical waves are more involved in inducing or suppressing precipitation than in maintaining it. This aligns with the results of the study by Ageet (2023) in East Africa.

## 7.4 Summary

Rainfall predictions provided by operational services are deemed unreliable in tropical Africa due to the dependence of NWP models on parameterized convection, which results in inaccurate forecasts of MCSs. Although convection-permitting models (e.g., Cafaro et al., 2021; Woodhams et al., 2018) may partially resolve these challenges, their adoption on a larger scale at storm-resolving resolutions is hindered by enormous costs and high energy requirements. From the findings outlined in Chapter 2, it is evident that data-driven short-term rainfall forecasts for tropical Africa can be made using predictive information contained in tropical waves. These forecasts are superior in accuracy and much cheaper to generate compared to NWP model-based forecasts, marking a significant advancement in the field.

By implementing the semi-automated strategy for selecting TW-based predictors, as described in Section 5.2.4, across the analysis domain, precise regions exhibiting the greatest activity for each tropical wave were identified. Presented below is a summary of the findings from this analysis: 1) The predictive information gained from the downstream grid point is more dominant (61% of all grid points) compared to that from the target grid point over the analysis region. 2) TD and Kelvin waves act as the top predictors over much of tropical Africa (48%), with TD being the dominant predictor over most of the western Sahel. 3) The MRG wave is more dominant over the oceanic region off the coast of Guinea, accounting for around 15% of all grid points. Interestingly, the EIG wave rarely emerges as the top predictor, accounting for only around 9% of all grid points and is mostly seen over the Eastern Sahel. 4) Although the IG1 waves indicate high levels at various grid points in the northern Sahel and in the oceanic region in the Gulf of Guinea, in reality, these signals may be associated with MCSs and fast-moving TDs. 5) Slow-moving waves such as ER and MJO surprisingly emerge as leading predictors over roughly 18% of all grid points, with adjacent areas often showing domination by fast-moving waves like TDs. This is likely indicative of the generation of faster-moving waves within the wave envelopes of MJO and ER.

We proceed to examine the modulation of daily rainfall anomalies by tropical wave phases by stratifying and compositing raw rainfall anomalies when they correspond to different phases. The findings from the analysis can be summarised as follows: 1) The phases of the TDs, MRG and the Kelvin waves have the most significant impact on daily rainfall anomalies. Clear and distinct characteristics (such as positive and negative anomalies) aligning well with the waves' crests and troughs are observable throughout all phases of these waves. 2) The EIG and IG1 waves have some activity within their respective regions and exhibit their characteristics to some extent but do not have a strong impact on rainfall modulation. 3) The rainfall anomaly fields are generally well-aligned with the phases in the case of MJO and ER waves, but their modulation is not particularly strong. These results are well in line with the research of Schlueter et al. (2019b,a).

We then move on to investigate the impact of wave phases on the skill of rainfall forecasts, both in the context of predictions of occurrence (at a threshold of 0.2 mm per day) and the amount of rainfall. To accomplish this task, we examine the area-averaged BSS (for precipitation occurrence) and CRPSS (for precipitation amount) of the ENS forecast against the EPC15 ensemble forecast (reference) computed within the above-defined regions during different tropical wave phases (1-8). The results obtained from this analysis are summarized as follows: 1) All waves, except for the EIG wave, have a significant impact on the BSS of the ENS forecast, shown by the distinct peaks during the wet phases. While the BSS of the ENS forecast dips during the wet phases of the IG1 wave, no apparent modulation is observed in the case of the EIG wave. 2) TD displays the greatest variation in BSS, achieving a maximum value of -0.03 during the peak wet Phase (5) and a minimum value of -0.22 during the peak dry Phase (1). 3)

Analysis of CRPSS modulation by the TW phases yields similar results to the BSS, but with slight variations; TDs, Kelvin MRG and ER waves show spikes in CRPSS during the wet phase(s), while a more significant (than BSS) drop in skill is observed during the wet phases of IG1. The deviation of CRPSS with the wave phases is generally less drastic than the BSS, indicating a stronger impact on inducing (or suppressing) precipitation than on maintaining it.

# 8. Statistical forecasting of 24-hour rainfall accumulation using tropical waves

Chapter 6 demonstrates that statistical forecasts excel over their NWP counterparts in predicting the likelihood of rainfall occurrence in tropical Africa, even after postprocessing the latter with a powerful tool such as IDR. Nonetheless, the prediction of precipitation occurrence alone is not the optimal approach. A statistical forecast of precipitation amounts would prove to be more versatile and ultimately more advantageous. However, utilising the approach employed in Chapter 6 to accomplish this may be sub-optimal as it fails to integrate the existing knowledge of coherent wave features directly. This is because a tool such as the CPA might not be capable of detecting non-linear interactions, primarily in areas where multiple waves can influence rainfall patterns. Chapter 5 presents a gradient-boosting method that not only overcomes this deficit but also semi-automates the process, thereby limiting the explicit need for human expertise in predictor selection. Chapter 7 illustrates how the phases of convectively coupled tropical waves affect daily rainfall anomalies but, more importantly, demonstrates the impact of tropical wave phases on the skill of ENS forecasts over tropical Africa, confirming the claims and findings of several earlier studies (e.g., Dias et al., 2018; Schlueter et al., 2019a,b). Building upon this foundation, the primary objective of this chapter is to develop statistical models trained exclusively on tropical-wave-based predictors to forecast the amount of daily precipitation in tropical Africa. To accomplish this, the present chapter will adhere to the following structure:

Due to the large-scale nature of TWs, their direct application as predictors at individual grid points may yield less than optimal results. Therefore, following the strategy followed in the previous chapter to identify the regions of highest wave activity, the local impact of TWs is calculated at every gridpoint through the computation of their local phases and amplitudes, as detailed by Section 5.2.3. The combined phase-adjusted wave amplitude (PWA) is then utilized to identify local predictors for every target grid point. Subsequently, we generate deterministic forecasts of rainfall amounts at an example grid point (13° N, 2° E) near Niamey in an operational setting with a 6-hour lead time using the gamma regression model and the CNN model described

in Chapter 5. We then transform them into probabilistic forecasts using the recently introduced EasyUQ (Walz et al., 2022, see section 4.4). Furthermore, to substantiate the importance of data-driven forecasts, we gauge the skill of the statistical forecasts against four benchmarks: EPC15-ensemble (**?**), EPC15-EasyUQ, ENS and ENS-Control EasyUQ forecasts, first over the example grid point near Niamey and then over the entire analysis domain. Lastly, the findings of this chapter are summarized in the Summary section.

## 8.1 Forecast performance of statistical models

As described, the purely data-driven models use the phase-adjusted wave amplitude as predictor variables to forecast daily rainfall accumulation with a 6-hour lead time. Figs. 8.1a (gamma regression) and 8.2a (CNN) show the distribution of observed (blue) and deterministic forecast (orange) of daily accumulated rainfall and are compared to assess the ability of the statistical models to learn the general distribution of daily rainfall accumulation. This is the primary evaluation of the statistical models' capabilities. The observed rainfall histogram indicates that zero rainfall occurred during a large number of timesteps ($\sim$ 500 out of 920), leading to a long right tail. The histogram of the gamma-regression model's deterministic forecast displays a peak at approximately 3mm and predicts minimum and maximum rainfall values of 0.36 mm and 451.28 mm, respectively. Despite this, the gamma-regression forecast does display a right tail, albeit less pronounced than the observation. These values reflect the model's limitation in predicting rainfall values approaching zero.

The CNN model's deterministic forecast produces a distribution that more closely matches the observed rainfall distribution. The CNN forecast contains more rainfall predictions of zero ($\sim$ 127 out of 920 timesteps), but still much lower than the observed rainfall. The CNN model also generates a forecast distribution with a long right tail and predicts a minimum and maximum rainfall of 0 mm and 52.42 mm, respectively. The deterministic forecast from the CNN model performs marginally better in this regard, as it can occasionally predict zero rainfall. It is worth noting that for the gamma regression forecast, we assume a gamma distribution, but for the CNN model, no distribution is assumed. The presented findings demonstrate that, although it is plausible for both models to forecast substantial amounts of rainfall, which may even exceed realistic levels (as in the case of the very unlikely daily rainfall amount of 451.22 mm), both models have acquired a reasonable understanding (the CNN better than the gamma regression) of the typical distribution of daily rainfall over the example grid point.

In comparison with the benchmark forecasts (ENS, ENS-Control EasyUQ, EPC15-ensemble and EPC15-EasyUQ forecasts, respectively) shown in Figs 8.3 –8.6 the following are evident: 1) Comparison of the panel (a) of the benchmark forecasts reveals that histograms of rainfall distribution for the EPC15 forecasts (Figs. 8.5a and 8.6a) are erroneous, as they rely only on climatological averages. 2) The ENS forecast and the

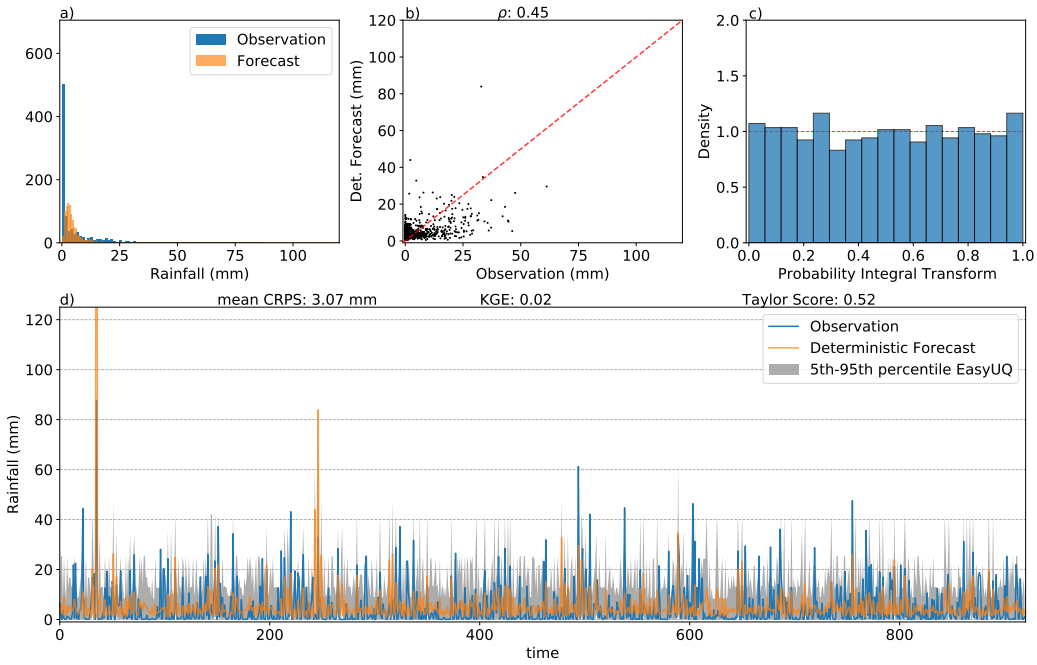



**Figure 8.1:** Illustration of the forecast of 24-hour rainfall accumulation at the gird point near Niamey(13° N, 2° E) via gamma-regression model using phase-adjusted wave amplitudes as predictors: a) Histogram showing the distribution of observed (blue) and forecasted (orange) rainfall; b) Observed and forecasted rainfall plotted against each other. This acts as a visual representation of the Pearson correlation ($\rho$) between the two quantities. The red dashed line along the diagonal indicates the line where $\rho = 1$, which is representative of the perfectly correlated forecast; c) The probability integral transform (PIT) histogram indicative of the calibration of the probabilistic forecast generated using EasyUQ. The red dashed line indicates a standard uniform distribution and acts as a reference to asses miscalibration; d) The time series displays observed (blue) and forecasted (orange) rainfall for all JAS seasons from 2007-2016. The grey shaded region is illustrative of the $5^{\text{th}}$ to the $95^{\text{th}}$ percentile range of the probabilistic forecast generated from the deterministic forecast based on the gamma regression model using EasyUQ.

ENS-control forecast (8.3a and 8.4a, respectively) have comparable distributions with the observation as they are generated by the ECMWF Integrated Forecasting System (IFS), an NWP model designed to simulate the atmosphere by solving prognostic equations that describe the atmosphere (and oceans) given the necessary boundary conditions and initial state.

Figs. 8.1b and 8.2b show the scatter between the observed and forecasted daily rainfall from the gamma regression and CNN models, respectively. The Pearson correlation coefficient ($\rho$) is shown in these graphical representations where a red dashed line along the diagonal refers to perfect correlation ($\rho = 1$). Like earlier, comparison with the benchmark forecasts (panel (b) of Figs 8.3 –8.6) reveals that the deterministic forecasts from the statistical models show much higher correlations with the observations compared to the benchmarks as shown in table 8.1.

As outlined in section 4.5, the Kling-Gupta Efficiency (KGE) and Taylor score, in addition to the Pearson correlation, are used to evaluate the deterministic forecasts

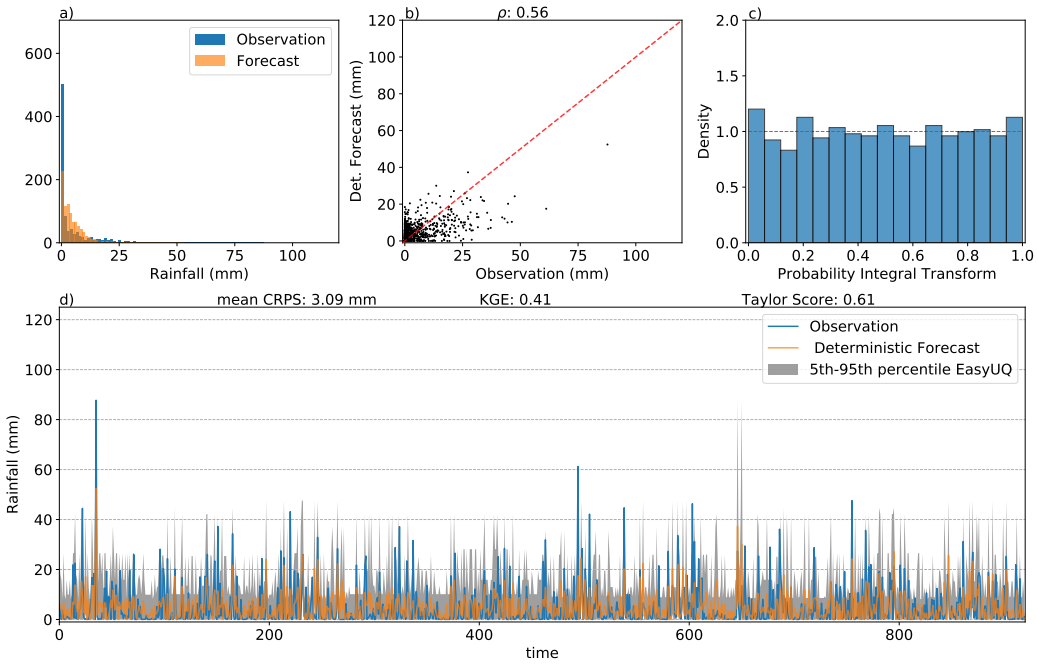



**Figure 8.2:** Same as 8.1 but for CNN model.

produced by the statistical models. Table 8.1 displays these metrics. The deterministic forecast generated by the CNN model exhibits the optimum values for all three metrics. This is due to the CNN model's design to maximise these metrics. Comparing the metric scores of all deterministic forecasts, it is evident that the ENS control forecast is the second-best (0.58) after the CNN forecast, surpassing the gamma regression forecast (0.52), with the exception of the Pearson correlation. This is especially apparent in the case of KGE. Both the Kling-Gupta Efficiency and Taylor Score assign high values to forecasts that replicate the general characteristics of the observations, including mean, variance, and correlation between forecast and observation. Although the gamma regression forecast displays a notable correlation with the observation, its KGE is suboptimal, likely due to the high standard deviation due to a few exceptionally high values. On the other hand, the ENS control forecast exhibits superior performance than gamma regression in this aspect, thereby achieving a comparatively higher KGE score. According to these metrics and CRPS (MAE in the case of deterministic forecasts) comparison in Table 8.2, the CNN deterministic forecast proves to be the best performing among other deterministic forecasts (with a CRPS of 4.98 mm).

It should be emphasized that even perfect values of the Pearson correlation, the KGE and the Taylor score do not guarantee a perfect forecast. A perfect forecast must not only simulate the magnitude of the observation but also do so at the right time. As an example, a comprehensive analysis of Figs. 8.1b and 8.2b demonstrates that although the CNN deterministic forecast outperforms the benchmark deterministic forecasts (Table 8.1), it still tends to produce outputs that are underestimated, as the majority of data points are below the diagonal. This reveals that despite providing reasonable

| Forecast | Pearson correlation | Kling-Gupta Efficiency | Taylor Score |
|---|---|---|---|
| Gamma regression | 0.45 | 0.02 | 0.52 |
| CNN | **0.56** | **0.41** | **0.61** |
| ENS-control | 0.24 | 0.17 | 0.58 |
| EPC15-ensemble | 0.18 | -0.16 | 0.07 |
| EPC15-EasyUQ | 0.18 | -0.16 | 0.07 |

**Table 8.1:** The Pearson correlation, Kling-Gupta Efficiency (KGE), and Taylor score are shown for the deterministic predictions of the statistical models and the benchmarks at the example grid point near Niamey. The highest scores are highlighted in bold. Higher values are better.

estimates of the daily rainfall distribution during JAS, both statistical models do not deliver timely and accurate forecast estimates. To address this issue, a range of studies (e.g., Gneiting, 2008; Gneiting and Katzfuss, 2014; Palmer, 2002) recommend the use of probabilistic forecasts, particularly when predicting synoptic-scale rainfall amounts. Probabilistic forecasts estimate uncertainties in future predictions and provide a level of confidence for each prediction (Gneiting and Katzfuss, 2014).

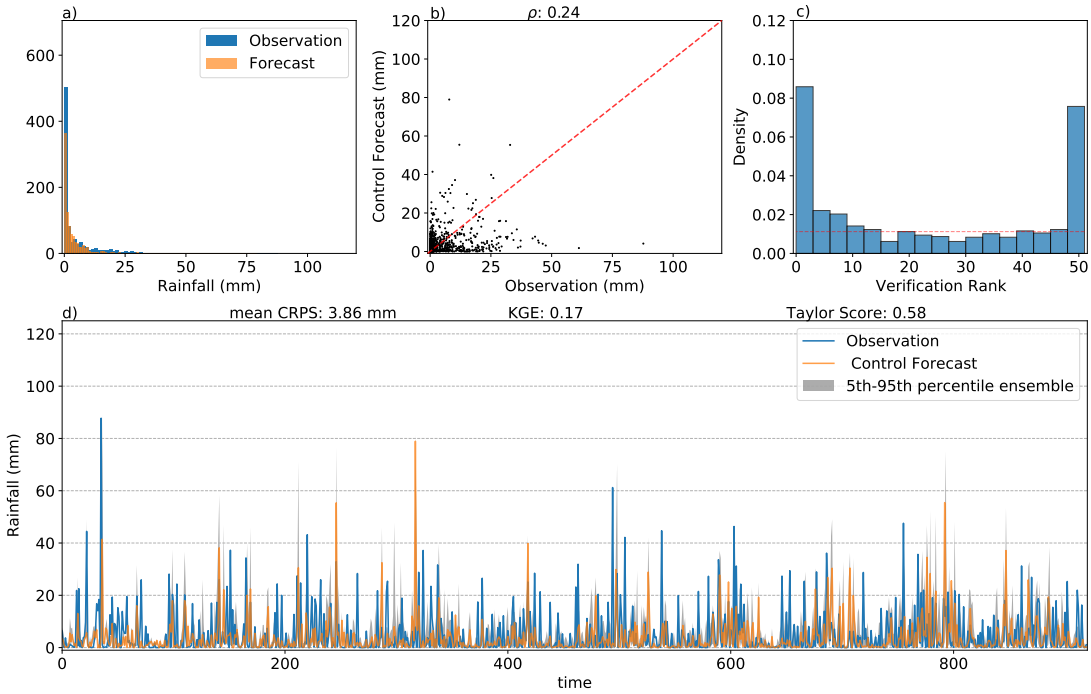



**Figure 8.3:** Same as 8.1 but for ENS forecast. c) Shows the histogram of verification ranks instead, and the red line indicates the median value of the ensemble forecast probability.

To accomplish this, we utilise EasyUQ, a newly developed tool that transforms deterministic forecasts created by statistical models into probabilistic forecasts, as detailed in Walz et al. (2022, refer to section 4.4 for more information). We additionally produce probabilistic forecasts derived from the benchmark deterministic forecasts in the same way to act as the new benchmarks.

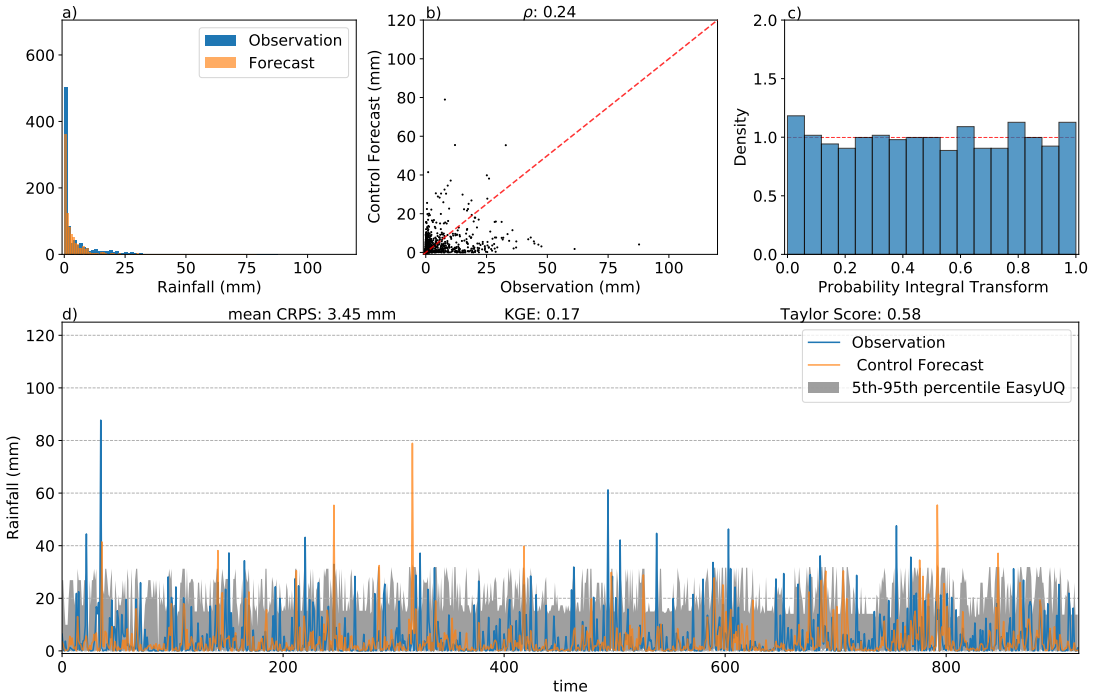



**Figure 8.4:** Same as 8.1 but for ENS forecast run through EasyUQ. Histogram showing the distribution of observed (blue) and (control) forecasted (orange) rainfall.

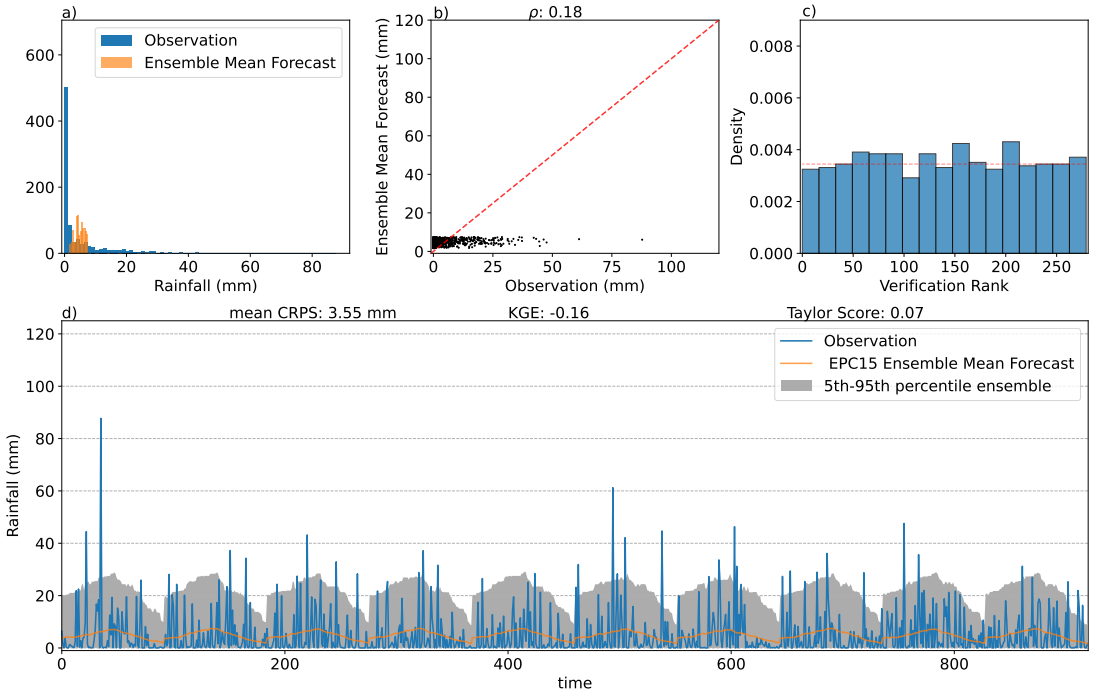



**Figure 8.5:** Same as 8.3 but for EPC15 ensemble forecast.

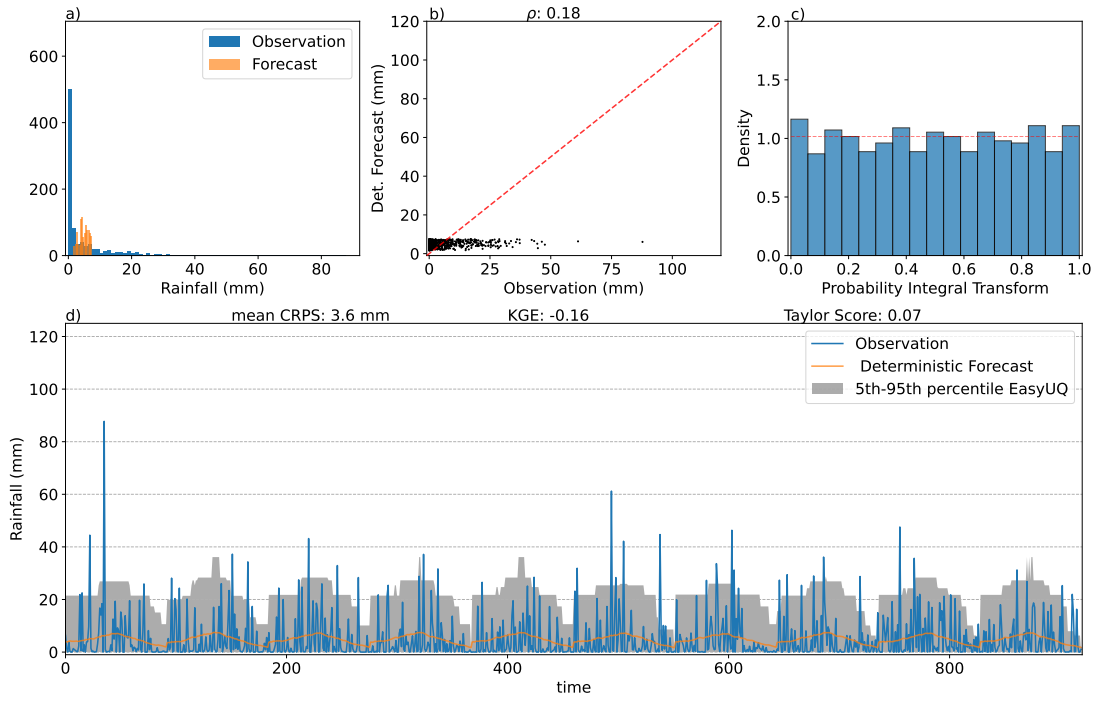



**Figure 8.6:** Same as 8.1 but for EPC15 run through EasyUQ. a) Histogram showing the distribution of observed (blue) and mean forecasted (orange) rainfall.

Panel (c) of Figs. 8.1-8.6 displays the probability integral transform (PIT) histograms (e.g., Gneiting et al., 2005; Dawid, 1984, see section 4.5.3.5 for details) of the probabilistic forecasts generated from the statistical model outputs (Figs. 8.1 and 8.2) and the benchmark deterministic forecasts (Figs. 8.3-8.6). PIT histograms are calculated using the probability integral transform theorem, which states that when the cumulative distribution function of any random variable is considered as another random variable, the resulting distribution of the new random variable is always standard uniform. In such cases, the PIT histograms of probability forecasts that follow a standard uniform distribution are considered calibrated, meaning that all probabilities should have the same density. When dealing with ensemble forecasts, a verification rank histogram is used (e.g., Anderson, 1996; Hamill and Colucci, 1997) instead, but the same result applies, where flat histograms signify calibrated forecasts, allowing direct comparison with PIT histograms. Hence, in the case of ENS and EPC15-ensemble forecasts, rank verification histograms are shown instead.

The PIT (verification rank) histograms of the probabilistic (ensemble) forecasts reveal the following: 1) Both the EPC15-ensemble (8.5c) and EPC15-EasyUQ (8.6c) forecasts exhibit good calibration, as expected. However, the EPC15-EasyUQ forecast appears slightly more accurate. This can be attributed to the fact that the EPC15 mean forecast is derived from the observed climatology and is, therefore, well-calibrated to begin with. Following this, EasyUQ transforms the forecast into a probabilistic one, which is likely to improve the calibration.; 2) Among the ENS (see Figure 8.3c) and the ENS-EasyUQ (see Figure 8.4c) forecasts, it is evident that the former is miscalibrated, evident from

the 'U-shaped' histogram, which is a well-known issue with the NWP ensemble forecasts in the tropics. This particular issue was discussed in detail in Chapter 6 concerning rainfall occurrence probabilities. Conversely, the ENS-EasyUQ forecast appears to be well-calibrated. This is analogous to the IDR-based statistical post-processing that improves the calibration of the ENS ensemble forecasts and hence demonstrates the versatility of IDR-based statistical tools; 3) It is noteworthy that the ENS-EasyUQ and EPC15-EasyUQ forecasts demonstrate comparable PIT histograms, signifying satisfactory calibration; 4) Both probabilistic forecasts based on statistical models appear well-calibrated as evident from the relatively flatter histograms, as expected for forecasts generated from observations. However, the CNN model-based forecast's PIT histogram (8.2c)) appears marginally more uniform than the gamma regression-based forecast (8.1c)), suggesting that the CNN forecast may be slightly better calibrated. This could be due to the CNN model's greater complexity compared to the gamma regression model, which may enable it to derive more intricate relationships from the predictor variables.

| Forecast | CRPS (mm) |
|---|---|
| Gamma regression deterministic | 5.18 |
| Gamma regression probabilistic | **3.07** |
| CNN deterministic | 4.68 |
| CNN probabilistic | **3.09** |
| ENS control | 5.11 |
| ENS ensemble | 3.86 |
| ENS-Control EasyUQ | 3.45 |
| EPC15-ensemble mean | 5.64 |
| EPC15-ensemble | 3.55 |
| EPC15-probabilistic | 3.60 |

**Table 8.2:** The Continuous Ranked Probability Score (CRPS) values are presented for the forecasts derived from statistical models as well as the benchmarks at the example grid point near Niamey. ENS-probabilistic and EPC15-probabilistic represent the probabilistic forecasts produced by EasyUQ from ENS control, and EPC15 mean forecasts, respectively. It should be noted that for deterministic forecasts (the control forecast in the case of ENS and the EPC15 ensemble mean forecast), MAE is displayed instead as CRPS reduces to MAE in these cases (see section 4.5.3.4). The lowest values are highlighted in bold.

Finally, the time series of the daily observed (blue) and the corresponding forecasted (orange) rainfall amounts at the example grid point near Niamey from the period 2007-2016 summer season are shown in the panel (d) of Figs. 8.1-8.6. The figures also show the $5^{\text{th}}$–$95^{\text{th}}$ percentile range (grey shading) of the probabilistic (ensemble) forecast computed using EasyUQ for the deterministic (ensemble) forecasts. Note that for both the EPC15-Ensemble forecast (seen in Fig. 8.5d) and the EPC15-EasyUQ forecast (seen in Fig. 8.6d), the deterministic forecasts are equivalent to the mean value of the distribution at each point in time.

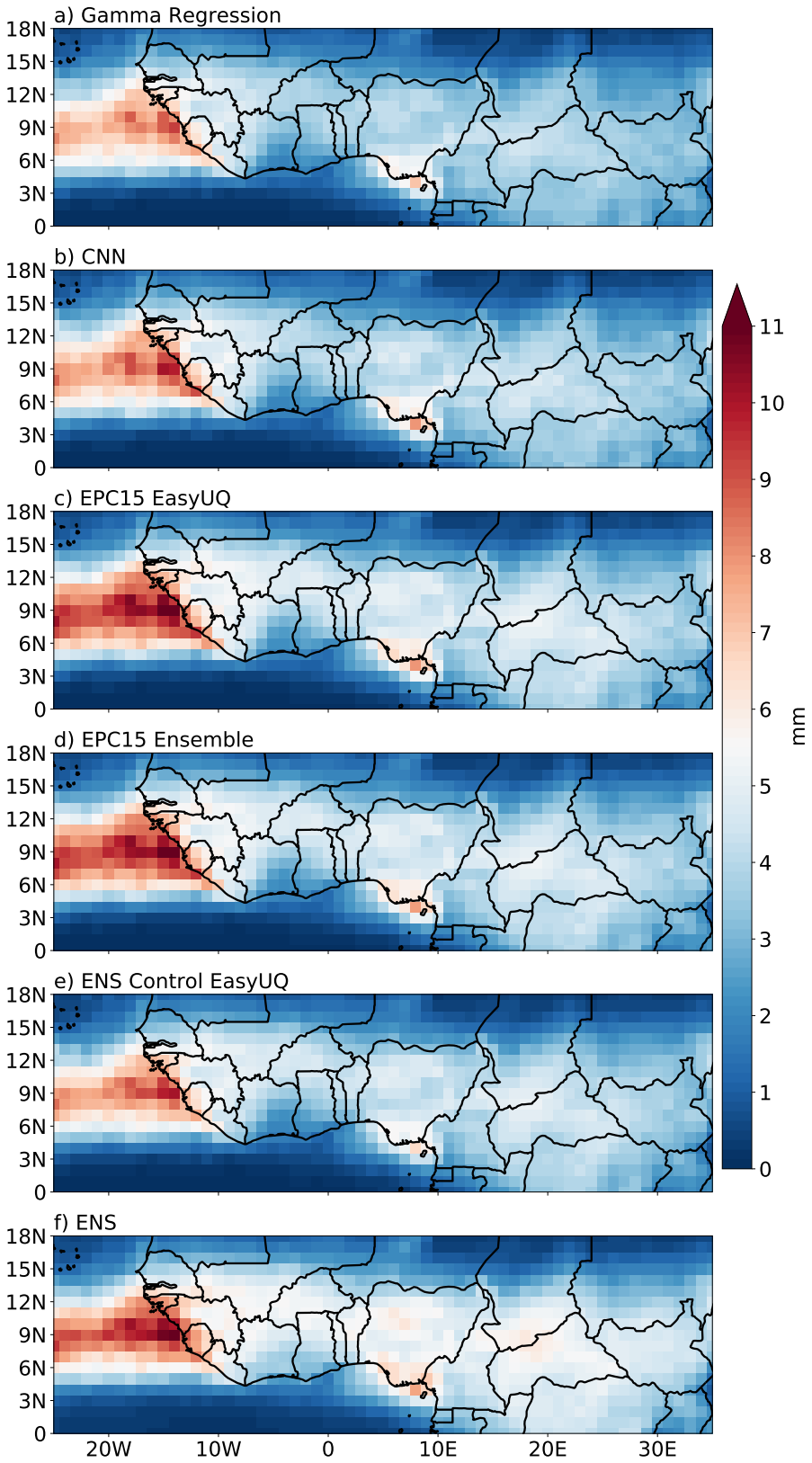



**Figure 8.7:** Continuous Ranked Probability Score (CRPS) for: a) deterministic forecast based on gamma regression, b) deterministic forecast based on CNN, c) mean forecast of EPC15, and e) ENS control forecast, all transformed into probabilistic forecasts via EasyUQ. Additionally, d) and f) show the CRPS for the EPC15 ensemble forecast and the ENS ensemble forecast, respectively.

From the analysis of the time series, the following conclusions can be drawn: 1) Apart from the identical deterministic forecasts provided by the EPC15-ensemble and the EPC15-EasyUQ models, their probabilistic forecasts also bear significant resemblance. Both models indicate similar CRPS values (3.55 mm and 3.60 mm, respectively), presented in Table 8.2, which clearly outperform the CRPS (5.64 mm) of the deterministic EPC15 forecast; 2) The probabilistic forecast using ENS-EasyUQ demonstrates a marked enhancement over the ENS ensemble forecast, with CRPS values of 3.45 mm compared to 3.86 mm (as displayed in Table. 8.2). This is interestingly in contrast to the results of Walz et al. (2022), who illustrated EasyUQ with WeatherBench temperature data, where the EasyUQ probabilistic forecast failed to outperform the CRPS of the ensemble forecast. This serves as an example of the reliability of EasyUQ in providing calibrated probabilistic forecasts of precipitation in the tropics. Nevertheless, as one would expect, the ENS ensemble forecast outperforms the ENS control forecast (MAE of 5.11 mm); 3) When assessing all variations of the EPC15 predictions against the ENS forecasts, it is worth noting that the EPC15 ensemble and probability forecasts excel, marginally under-performing only with respect to the ENS-EasyUQ forecast, once again illustrating the poor performance of NWP model-based precipitation forecasts over tropical Africa; 4) the deterministic forecasts from both statistical models demonstrate comparable performance, with a MAE of 5.18 mm and 4.68 mm for the gamma regression and the CNN models respectively; 5) Demonstrating the robustness of the combined use of the statistical models and EasyUQ, the probabilistic forecasts based on the statistical model outputs, significantly outperform all other forecasts in the gridpoint near Niamey, with CRPS values of 3.07 mm and 3.14 mm for the gamma regression and CNN models, respectively.

We repeat this analysis at all grid points in the analysis domain shown as the nested core area in Fig. 5.2. Forecasts will only be issued in this area due to the considerable computational load compared to the analyses of Chapter 6. To this end, we calculate the CRPS across all grid points in the analysis domain for each of the statistical forecasts to facilitate comparisons of their performances with the benchmarks, as depicted in Fig. 8.7.

The following interpretations can be made from the CRPS of the forecasts in the analysis domain: 1) All forecasts show the highest CRPS over the western coasts of West Africa (Guinea, Sierra Leone, Liberia, etc.) and along the coasts of Cameroon and Nigeria. This outcome is unsurprising, as these regions experience the most rainfall in the JAS season and are located along the African rain belt. When examining the forecasts, it becomes clear that the ENS forecast (illustrated in Fig. 8.7f) shows the highest CRPS (above 15 mm) over these regions, while the two statistical models (Fig. 8.7a, b) show the lowest values ($\sim$ 8 mm); 2) The regions with the lowest CRPS values are located over the oceanic region along the Gulf of Guinea, the northern equatorial Atlantic (the coast of Senegal) and the northern Sahel. Notably, unlike in point 1, there are no significant discrepancies in the CRPS values among various forecasts in these

regions, primarily because they are among the driest regions receiving less than $\sim 50$ mm per month during the summer season; 3) Over the rainbelt region in continental tropical Africa, which includes the Sahel, the Guinea coast, and central Africa, again the ENS forecast indicates CRPSs with the highest values, reaching up to 7.5 mm. In comparison, the two statistical models demonstrate the lowest CRPS values, reaching only 5.5 mm; 4) All forecasts display a 'kink' of low CRPSs along the border between Cote d'Ivoire and Ghana, signifying a realm of higher predictability between the high CRPS regions that have been discussed in point 1. It is worth noting that this pattern is less evident in the ENS forecast, and the difference in CRPS between the two statistical forecasts and the ENS forecast is significant (around 3 mm) at this location; 5) All benchmark forecasts, except for the ENS ensemble forecast (shown in Figs 8.7c, d, and e), display similar CRPSs, indicating comparable predictability. However, upon closer examination, it becomes apparent that the EPC15-ensemble forecast (shown in Fig 8.7c) exhibits slight superiority (around 0.4 mm ) in the rain belt region of continental Africa, while the probabilistic forecast generated from the ENS-control forecast shows a slight improvement (just over 0.4 mm) over the high CRPS regions mentioned in point 1 (not shown).

To enable a direct and objective comparison between the statistical and benchmark forecasts, we calculate the Continuous Ranked Probability Skill Score (CRPSS; refer to section 4.5.3.4 for more information). Figs. 8.8 and 8.9 present the CRPSS for probabilistic forecasts based on gamma regression and CNN deterministic forecasts against the benchmark forecasts, respectively. Overall, both statistical forecasts tend to outperform the benchmark forecasts, particularly over regions that coincide with the African rainbelt over the land, such as the Sahel, the majority of the Atlantic Ocean situated to the west of the analysis domain and central Africa, with some marked differences between different benchmarks. The main inferences that can be drawn are as follows: 1) Over common grid points, where both statistical forecasts significantly outperform the benchmarks, gamma regression forecast (Fig. 8.8) tends to perform better, indicated by the darker blue shading. 2) Both statistical forecasts exhibit superior performances compared to the raw ENS forecast (panel (d) of Figs. 8.8 and 8.9 for gamma regression and CNN, respectively), across most areas of analysis domain. These findings indicate a substantial miscalibration in the ENS ensemble forecast for rainfall in tropical Africa, consistent with the findings of Chapter 6. It can also be seen that the ENS forecast exhibits better performance at some grid points in the north-western (southwest coast of Mauritania and Senegal) and north-eastern (central Chad and Sudan) regions. A potential cause for this may be that these areas are far enough from the equator to experience extra-tropical (see Section 2.1.4.2 for details) influences, which could improve the performance of the ENS forecast. Additionally, the statistical forecasts are incapable of incorporating extra-tropical effects, as they are trained solely on data based on TWs limited to three zonally aligned grid boxes. However, as indicated by the hatching based on the Diebold-Mariano test for equal predictive power, the ENS forecast,

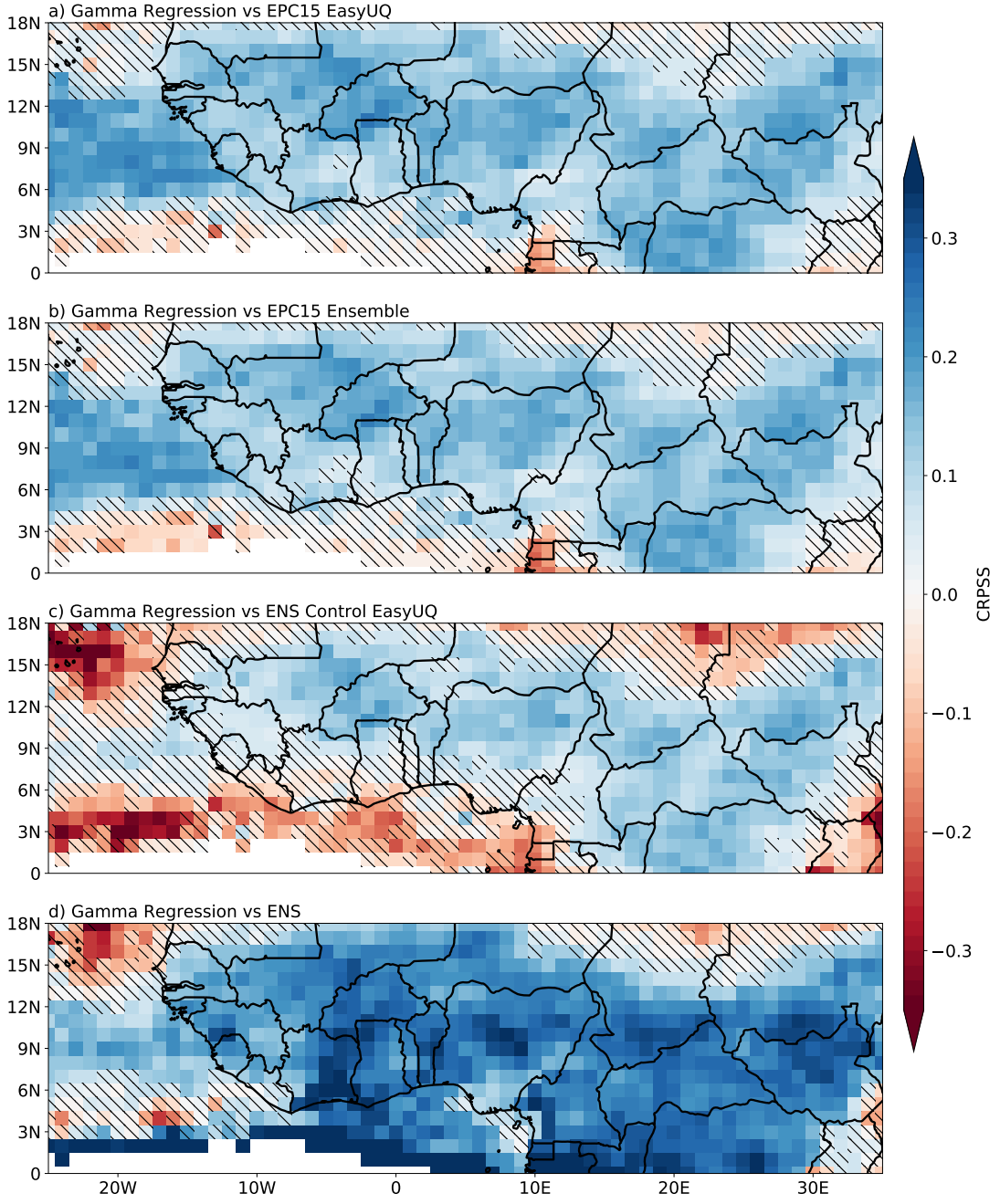



**Figure 8.8:** The Continuous Ranked Probability Skill Score (CRPSS) of gamma regression probabilistic forecast relative to a) the EPC15-probabilistic forecast, b) the EPC15-ensemble forecast, c) the ENS-control based probabilistic forecast and d) the ENS forecast. Hatching indicates grid boxes where the difference in the mean CRPS scores fail to be statistically significant at a level of 0.05 according to the Benjamini-Hochberg corrected Diebold-Mariano test for equal predictive ability. Grid points with less than 5mm of average monthly rainfall are left blank.

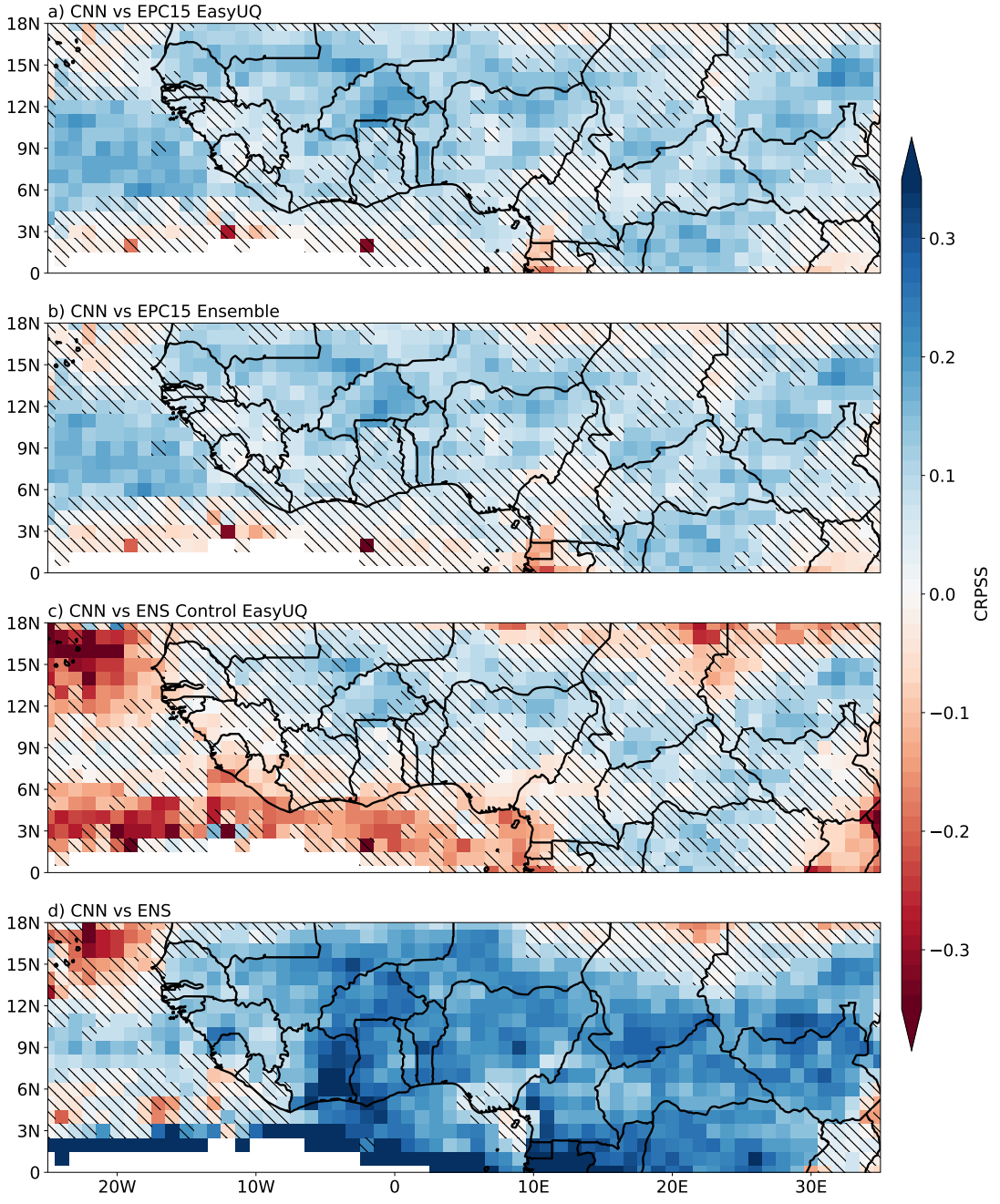



**Figure 8.9:** Same as Fig. 8.8 but for CNN probabilistic forecast.

while potentially better, does not significantly outperform the statistical forecasts over most of these regions.

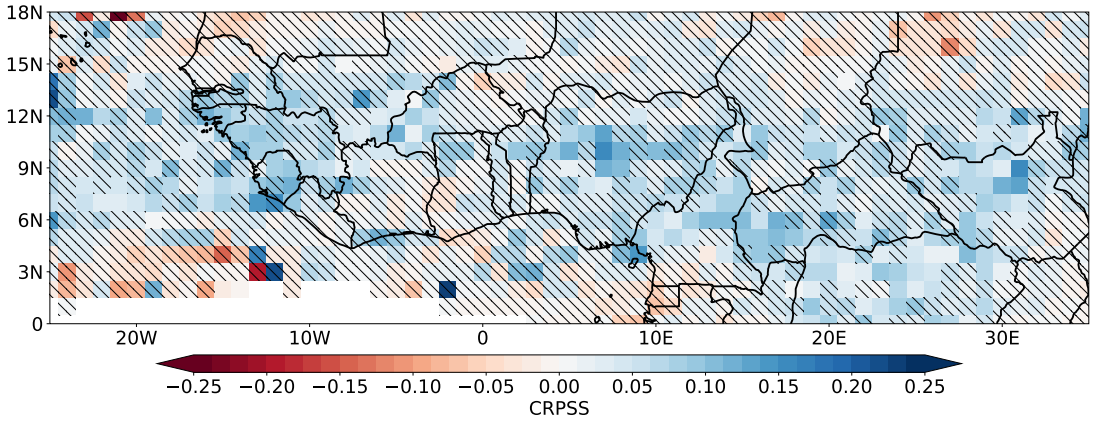



**Figure 8.10:** CRPSS of the gamma regression-based probabilistic forecast relative to the CNN-based probabilistic forecast. Note that the scale of the colorbar is adjusted for better clarity. Other details remain the same as in Fig. 8.8.

3) The CRPSS of the statistical forecasts relative to the ENS control EasyUQ forecast (panel (c) of Figs. 8.8 and 8.9 for gamma regression and CNN respectively) reveals a mixed result. The statistical forecasts exhibit superior performance in the land regions that align with the African rainbelt; however, the CNN forecast only barely outperforms the benchmark. The gamma regression forecast performs better over most of the land regions and parts of the oceanic region off the coast of Guinea, coinciding with the rainbelt. Nevertheless, there are minor variations in skill depending on the location. The best performance (above 0.15) is observed over the land, in southern Mali, Burkina Faso and southern Niger in the western Sahel, northern DRC and north-eastern Congo in central Africa and parts of southern Sudan (not South Sudan) in the eastern Sahel (Fig. 8.8c). In contrast, the ENS-control-EasyUQ is better over most of the northern grid points in the analysis area (especially the offshore regions of Senegal in the west and central Chad, central Niger and Sudan in the east), the Gulf of Guinea, the equatorial Atlantic, and parts of Uganda. The improved performance of the ENS-control EasyUQ forecast in the northern extent of the analysis region can be attributed to the same reasons outlined in point 2. However, the lower skill of the statistical models over the equatorial Atlantic may be due to the presence of numerous zero precipitation timesteps, as rainfall during JAS is infrequent in this region. The superior capability of ENS in predicting orography-driven precipitation is likely responsible for its improved performance in Uganda (e.g., Lafore et al., 2017; Vogel et al., 2020). Nevertheless, for the majority of these grid boxes, the ENS control EasyUQ forecast is not significantly more skilful than the statistical forecasts (especially the gamma regression-based forecast), as indicated by the hatching. Notable exceptions to this include the relatively drier oceanic regions of the coast of Senegal, the equatorial Atlantic and parts of the northeastern Sahel.

4) When comparing the CRPSS of the statistical forecasts relative to the EPC15-based probabilistic forecasts (presented in panels (a) and (b) of Figures 8.8 and 8.9 respectively, for both gamma regression and CNN models), no significant differences in the results were found in comparison to those discussed in point 3. However, it is worth noting that the EPC15 ensemble forecast marginally (but significantly) outperforms the gamma regression-based forecast over parts of the equatorial Atlantic region (notice the unhatched grid points in Fig. 8.8b). However, the CNN-based forecast performs better in this regard (notice the hatched grid points over the same region in Fig. 8.9b). Also, both versions of EPC15-based forecasts outperform both statistical forecasts over Equatorial Guinea. The reasons for this are currently unclear.

5) Comparing Fig. 8.8 with Fig. 8.9, it becomes clear that the gamma regression-based probabilistic forecast is considerably more adept than the CNN-based one over almost all regions. This is indicated by the darker and more blue-shaded grid points, as well as fewer hatched grid points. To illustrate this more clearly, we refer to Fig. 8.10, which displays the CRPSS of the gamma-regression-based forecast compared to the CNN-based forecast. Although supporting the aforementioned argument, Fig. 8.10 also reveals that there is no significant difference between the statistical predictions (indicated by hatching), except for a few grid boxes situated in the centre of the rainbelt.

## 8.2 Summary

Several studies (e.g., Schlueter et al., 2019a,b) have investigated the potential of using tropical waves as predictors for synoptic-mesoscale rainfall forecasting over the tropics. Thus, a data-driven approach for forecasting daily rainfall accumulation per grid point with a six-hour lead time is introduced to address this gap.

We utilise the local impact (i.e. at a grid point) of tropical waves, computed using their local amplitudes and phases as detailed in Section 5.2.3. We proceed to train two statistical models, specifically a gamma regression model and a convolutional neural network model, by utilizing the PWAs to generate deterministic forecasts of daily accumulated rainfall for the summer seasons (JAS) of 2007-2016. The deterministic forecasts are then transformed into probabilistic forecasts via the recently introduced EasyUQ (Walz et al., 2022) and compared with four benchmarks: the EPC15 ensemble forecast, the EPC15-EasyUQ forecast, the ENS raw forecast and the ENS-control EasyUQ forecast. The first and third benchmarks are ensemble forecasts, while the second and fourth are probabilistic forecasts produced by transforming the EPC15 mean forecast and the ENS control forecast, respectively, via EasyUQ. The forecast analysis summary is as follows: 1) The regions with the highest rainfall, which are located along the coasts of Guinea, Sierra Leone and Liberia, and a small region over the coast of Cameroon, pose a challenge for all the forecasts. 2) The probabilistic forecasts derived from the statistical model outputs outperform the ENS raw forecast over the land region, indicating serious miscalibration in the ENS forecast. 3) This problem is largely

resolved when producing probabilistic forecasts via EasyUQ using the ENS-control forecast. Consequently, the forecast skill comparison results in a mixed outcome overall. While the probabilistic forecasts based on the gamma regression model outperform even the calibrated ENS-control-based probabilistic forecast over the rain belt region, the latter performs better over regions that receive very low rainfall climatologically (such as the Atlantic Ocean off the coast of Senegal, the northeastern Sahel, some grid points over Uganda, and most of the grid points over the oceanic region along the equator) However, only a small fraction of these regions show significant differences, not enough to expect a consistent performance gain over the region. The forecast based on the CNN model, however, performs worse than the gamma regression model and only marginally outperforms the ENS-Control EasyUQ forecast over the land. 4) The comparison between the statistical forecasts and the EPC15-based forecasts yields comparable results. Nevertheless, the EPC15 ensemble forecast exhibits superior performance to the gamma regression forecast across the equatorial Atlantic and, notably, across certain regions of Equatorial Guinea and Gabon. However, it falls short of significantly outperforming the CNN forecast over these regions. 5) A comparison of the statistical models indicates that the gamma-regression-based forecast noticeably performs better than the CNN-based forecast in most of the rainbelt region. However, the latter shows improved performance over regions with very low rainfall, such as the northern Sahel and the Equatorial Atlantic Ocean, as well as over Equatorial Guinea and Gabon. Despite this, neither model is seen to possess a significant advantage over the other within the analysis domain, except over certain grid points along the regions coinciding with the rainbelt.

In view of these findings, it is important to emphasise that the subpar performance of the ENS forecast over Tropical Africa is attributed to its poor calibration. However, these are convection-parameterised forecasts, and convection-permitting forecasts may perform better in principle (e.g., Woodhams et al., 2018; Cafaro et al., 2021). But they will remain several times more expensive to produce than their convection-parameterized counterpart in the near future. A more cost-effective option is to employ tools such as EasyUQ to generate probabilistic forecasts from the deterministic output of an NWP model. This study shows that such methods greatly improve the skill of the ENS forecast over tropical Africa by producing calibrated forecasts but at a fraction of the cost of running ensemble forecasts. However, an even better option based on the findings of this chapter is to employ probabilistic forecasts generated from simple models like the gamma regression model in combination with EasyUQ. Except for specific regions, the gamma regression-based forecast exhibits superior performance even compared to a more sophisticated CNN-based probabilistic forecast over most of the analysis domain, likely due to limited data used in training. Although deep-learning models like CNN can exploit massive amounts of data and identify intricate patterns, simpler models like the gamma regression model employed in this study may produce better results with limited data.

# 9. Conclusions

Rainfall prediction in tropical Africa is very challenging. The challenges stem from various factors, including the inherent characteristics of convection-driving processes that modern-day numerical weather prediction (NWP) models struggle to resolve, as well as the reduced intrinsic predictability in the tropics. This lack of predictability leads to rapid error growth, which decreases the accuracy of NWP model-based forecasts. As a result, these forecasts are often only as effective as simple climatology-based forecasts at best, even after statistical correction for biases and dispersion errors.

Statistical models offer an alternative approach to weather forecasting. Only limited by the quality of the predictors and the specific nature of the model utilised, statistical models are able to learn the coherent relationship between rainfall and other meteorological variables to issue synoptic-scale rainfall forecasts over tropical Africa that outperform their NWP counterparts. An added advantage is that this is achieved at a fraction of the cost required to run NWP models with numerous ensembles. A recent investigation conducted by Vogel et al. (2021) has demonstrated that employing a straightforward Logistic regression model trained on prior rainfall observations can exceed the performance of the NWP-based ensemble forecast in terms of detecting daily precipitation over tropical Africa. Motivated by this, the primary objective of this thesis was developing statistical models using observation-based predictors for issuing forecasts of synoptic-scale rainfall over tropical Africa. Subsequently, the predictive ability of these models was evaluated systematically using a wide range of metrics and benchmarks. Where possible, the forecast results were substantiated by meteorological evidence.

Motivated by this prospect, the logistic regression (Logistic) model, developed in Chapter 6, was trained on historical rainfall data based on Global Precipitation Measurement Integrated Multi-satellitE Retrievals (GPM-IMERG) to predict the probability of daily rainfall occurrence in tropical Africa. Through a systematic analysis of the results, substantiated with meteorological reasoning, answers to the following research questions presented in Chapter 3 were obtained:

**RQ1a How does the skill of the Logistic forecast compare to that of the climatology and the NWP-based forecast (both raw and postprocessed)?**

The Logistic forecast excels primarily in areas along the fringes of the climatological rainbelt and the Gulf of Guinea, where the model identifies coherently propagating precipitation patterns influenced by dominant synoptic-scale drivers such as the AEWs and the westward propagating cyclonic vortices.

All forecasts struggle in regions that correspond with the climatological rainbelt, where precipitation is primarily influenced by extremely stochastic processes. A direct comparison of the Logistic forecast with the climatological forecast and the NWP-based ensemble forecast reveals that the Logistic forecast performs better nearly everywhere throughout the analysis domain. Although statistical postprocessing enhances the skill of the ensemble forecast, the gain in performance does not achieve statistical significance. With only marginal variations, these findings persist across other seasons as well.

**RQ1b How much predictive value do lags of one, two and three days provide?**

Most of the predictive information is contained in the one-day lag, especially over the Gulf of Guinea. Incorporating predictive information from lags of two and three days incrementally improves the skill of the Logistic forecast, according to our analysis. However, the gain in performance is marginal.

Information about the precipitation patterns from lags of two and three days, however, contributes substantially to providing meteorological reasoning for the Logistic forecast skill, especially in terms of *coh*.

**RQ1c How can the performance of the Logistic forecast be explained using coherent propagation of rainfall patterns and meteorological variables like Convective Available Potential Energy (CAPE), shear and Convective Instability (CIN)?**

Large values of *coh* indicate coherent propagation of rainfall patterns over three days. Such patterns usually correspond with high values of CAPE, CIN, and shear, requiring the presence of relatively strong triggers like AEWs and MCSs to generate precipitation. The superior skill of the Logistic forecast along the edges of the rain belt can be attributed to these factors.

In regions that coincide with the rainbelt, the values of *coh* are low, indicating a lack of coherently propagating rainfall patterns. When combined with high values of CAPE and low values of CIN, this indicates a high level of stochasticity. Consequently, the performance of the Logistic forecast does not show significant skill gain in such regions.

The findings of Chapter 6 indicate that statistical models leveraging the predictive information contained in coherently propagating precipitation patterns can outperform NWP model-based forecasts. As discussed in Chapter 2, the use of tropical waves as potential predictor variables is a possibility worth considering since they significantly contribute to the modulation of rainfall in tropical Africa, also detailed in several studies (Dias et al., 2018; Schlueter et al., 2019a,b). However, it is crucial to examine and measure the influence of tropical waves on rainfall modulation before proceeding. Additionally, it is worth investigating if this impact extends to the modulation of NWP model-based forecast skill.

Motivated by this, in Chapter 7, we initially identified the regions impacted by predictor variables based on tropical waves. Subsequently, we examined how tropical waves affect the modulation of daily rainfall anomalies in these areas. Finally, we analysed the influence of tropical wave phases on the skill of the ENS ensemble forecast relative to the EPC15 forecast to address the research questions established earlier in Chapter 3

**RQ2a How do the predictive abilities of various tropical waves vary across different regions of tropical Africa?**

The regions where different TW-based predictors have the greatest impact are shown in the table below:

| Wave | Regions |
|---|---|
| EIG | Central-northern Sahel (downstream) |
| ER | Gulf of Guinea, Atlantic Ocean (downstream) |
| IG1 | Sahel, Gabon–coastal regions (downstream) |
| IG1 | Guinea Coast (upstream) |
| Kelvin | Eastern Sahel, central Africa (downstream) |
| MJO | – |
| MRG | Congo Basin, central Sahel, Atlantic Ocean (downstream) |
| TD | Entire analysis region (downstream) |

Some points to consider: 1) The majority of predictive skill of statistical forecasts comes from predictors located downstream of the target grid point. IG1 is the only wave mode that shows predictive skill in upstream gridpoints. 2) While TD is an important predictor throughout most of the analysis domain, only those located in and near the Sahel region

may be considered as AEWs. 3) In comparison to other wave modes, MJO is not a significant predictor for six-hour ahead daily rainfall forecasts.

**RQ2b To what extent do the phases of tropical waves impact the skill of predicting precipitation occurrence and amounts?**

Our analysis indicates that the phases of TDs, MJO, Kelvin, MRG and ER waves have a significant impact on the skill of the NWP model-based forecast in predicting rainfall occurrence. This is demonstrated by clear peaks in the BSS during the wet phases. While the BSS of the ensemble forecast (relative to the climatology-based forecast) shows a slight dip during the wet phase of the IG1 wave, no significant modulation is seen during the phases of the EIG wave.

However, it is noteworthy that even at the highest value, the BSS of the ensemble forecast remains negative, highlighting the overall unsatisfactory performance when compared to the climatology-based forecast. Similar patterns are noticeable in the CRPSS as well, but they are less pronounced, suggesting that the phases of tropical waves have a greater impact on the occurrence of rainfall than the amount of rainfall.

As potential enhancements to the statistical method presented and discussed in Chapter 6, the procedures detailed in Chapter 5 include the development of a semi–automated feature selection procedure based on gradient-boosting and statistical models based on gamma regression and convolutional neural network (CNN). The models forecast daily rainfall amounts, providing a more beneficial prediction than probabilities of rainfall occurrence. Additionally, we also used the recently developed EasyUQ (Walz et al., 2022) to transform the deterministic outputs of the above-mentioned statistical models into probabilistic predictions. Additionally, the models developed in Chapter 8 were trained using predictors derived from convectively coupled tropical waves, following the conclusions drawn in Chapter 7. Two climatology-based forecasts and two NWP model-based forecasts (one ensemble and one probabilistic) were used as benchmarks to systematically evaluate the skill of the above-mentioned statistical forecasts. In light of these analyses, Chapter 8 addresses the research questions raised earlier in Chapter 3:

**RQ3a How do the skills of gamma regression and the CNN models measure up against those of the forecasts based on climatology and the NWP model-based forecasts (both ensemble and probabilistic)?**

Both gamma regression and CNN forecasts demonstrate good performance across the entire analysis domain, suggesting that tropical waves serve as valuable predictors for forecasting daily accumulated rainfall. All forecasts encounter difficulties over regions near the Guinea highlands and the Cameroon highlands, where the highest quantities of rainfall occur during the summer season. Nonetheless, the gamma regression and CNN forecasts exhibit lower CRPS values, indicating better overall performance.

Direct comparisons of the statistical and forecasts based on climatology and the NWP model indicate that both statistical forecasts demonstrate significantly superior performance compared to the climatology benchmarks and the ensemble forecast across most regions (particularly over rainbelt regions). Comparison with the NWP model-based probabilistic forecast indicates that both statistical forecasts display superior performances over the rainbelt region.

The NWP model-based probabilistic forecast exhibits better performance in certain areas of the northern Sahel and oceanic regions. Possible causes of this phenomenon over the Sahel may be tropical-extratropical interactions. The statistical models trained on convectively coupled tropical waves (TWs) may experience difficulties in accurately predicting extremely arid conditions. Such limitations may be accountable for their unsatisfactory performance over oceanic regions.

**RQ3b How does the skill of gamma regression forecast compare against that of CNN forecast?**

A comparison between the skills of the statistical models indicates that the forecast from the gamma regression model performs better than that from the CNN model in most of the rainbelt region. However, the converse is found to be true in regions with low rainfall, such as the northern Sahel and the equatorial Atlantic Ocean. One possible explanation may be that the CNN models can efficiently handle a large volume of zero precipitation events in the data, while a relatively simpler gamma regression forecast cannot, even after transformation into a probabilistic forecast. However, neither model displays a noteworthy advantage over the other in the analysis domain.

This thesis presents a thorough demonstration of the implementation of statistical forecasts that are trained solely on past observations for the prediction of synoptic-scale rainfall across tropical Africa. The findings not only emphasised the benefits of employing such strategies but, more importantly, offered detailed discussions of the potential meteorological factors contributing to the performance enhancements. To the best of our knowledge, this thesis presents the first in-depth account of the use of convectively coupled tropical waves in conjunction with machine learning models to produce daily rainfall forecasts over tropical Africa. By comparing a diverse selection of benchmark forecasts, including some solely based on NWP models, others relying on calibrated NWP model outputs, and some using basic climatology, this study highlights the advantages and disadvantages of statistical forecasting for the region of tropical Africa. Although some progress has been made through the implementation of our strategies, we acknowledge that predicting the precipitation and, in general, the weather in tropical Africa remains a challenging endeavour. We aspire that the findings presented in this thesis will aid in advancing this endeavour.

# 10. Outlook

With a view toward operational forecasting, we recommend that operational services consider the use of statistical post-processing techniques that improve the skill of the raw forecasts by proper calibration (Hewson and Pillosu, 2021; Henzi et al., 2021). Another alternative is to produce probabilistic forecasts by transforming the control forecasts produced by operational services, using tools such as EasyUQ (Walz et al., 2022), as they are readily calibrated and orders of magnitude cheaper than running ensemble (CP) forecasts. Climatological forecasts such as the EPC forecast offer very cheap alternatives if such tools are not available. However, our key recommendation is to utilise data-driven techniques, including the statistical forecasts that have been analysed, along with tools like EasyUQ, to obtain probabilistic precipitation forecasts for northern tropical Africa. This approach offers the additional advantage of being much more economically feasible than NWP model-based forecasts.

As possible enhancements to forecasting strategies for rainfall in tropical Africa, the following measures could be taken into account:

NWP forecasts from various operational centres could serve as benchmarks. This research focused solely on ECMWF's operational precipitation ensemble forecasts. Nevertheless, a multi-model analysis could uncover fresh insights that might aid in refining the benchmarking methodology. Another approach is to integrate data from adjacent grid points when postprocessing the NWP forecasts, which could enhance the overall skill.

The present thesis employed data sets with a resolution of $1^\circ \times 1^\circ$ and daily temporal scales. Though it served the purpose of reducing computational burden without losing the ability to detect synoptic-scale features like tropical waves, future studies can consider enhancing resolutions in both the spatial and temporal domains to account for localized variations. The predictive method employed in Chapter 8 could also be adjusted to investigate the feasibility of integrating predictors from additional grid points and perhaps even more meteorological variables. The primary aim of this dissertation was to develop statistical forecasting methods for predicting rainfall at daily scales. However, adapting these strategies and including longer lead times would be more advantageous. Including data from numerical weather prediction models in data-driven approaches (hybrid modelling) may also be considered, as demonstrated by studies such as Maier-Gerber (2023), to enhance the skill of the forecasts.

Unlike the predictor selection strategies employed in this study, which are only partially automated, more advanced techniques with complete automation capability could replace them. In line with the theme of this thesis, specifically presenting meteorological evidence to corroborate the outcomes of statistical forecasts, more advanced tools such as SHAP (SHapley Additive exPlanations; Ribeiro et al., 2016; Štrumbelj and Kononenko, 2014) could be utilised for this objective. While modern neural networks have the capability to learn patterns from data without the need for variable selection by humans, it remains to be seen whether they can surpass simpler data-driven or hybrid (statistical-dynamical) methods in areas such as northern tropical Africa.

Finally, one might employ advanced techniques of machine learning, such as graph neural networks, to develop more sophisticated data-driven approaches, as demonstrated impressively by Lam et al. (2022), Bi et al. (2023), Nguyen et al. (2023), and Chen et al. (2023). Whilst artificial intelligence (AI) based data-driven methods such as these show immense potential for many aspects of forecasting, they still face hurdles in forecasting precipitation. The approaches demonstrated in this thesis may help advance AI-based methods in this direction as our rather simple data-driven Logistic forecast and Gamma regression forecast combine expertise from data science and atmospheric science to leverage insights about spatio-temporal patterns of precipitation over northern tropical Africa and adjacent regions.

# Bibliography


Abbe, C., 1901: The physical basis of long-range weather forecasts. *Monthly Weather Review*, **29 (12)**, 551–561.

Ageet, S., 2023: Predictability of rainfall over equatorial east africa from daily to sub-monthly time scales using multi-model ensemble forecasts. Ph.D. thesis, Dissertation, Karlsruhe, Karlsruher Institut für Technologie (KIT), 2023.

Agustí-Panareda, A., A. Beljaars, C. Cardinali, I. Genkova, and C. Thorncroft, 2010: Impacts of assimilating AMMA soundings on ECMWF analyses and forecasts. *Weather and Forecasting*, **25 (4)**, 1142–1160.

Albright, M. D., E. E. Recker, R. J. Reed, and R. Dang, 1985: The diurnal variation of deep convection and inferred precipitation in the central tropical Pacific during January–February 1979. *Monthly Weather Review*, **113 (10)**, 1663–1680.

Anderson, J. L., 1996: A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *Journal of climate*, **9 (7)**, 1518–1530.

Ayesiga, G., C. E. Holloway, C. J. Williams, G.-Y. Yang, and S. Ferrett, 2021: The observed synoptic scale precipitation relationship between western equatorial africa and eastern equatorial africa. *International Journal of Climatology*, **41**, E582–E601.

Bai, S., J. Z. Kolter, and V. Koltun, 2018: An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525 (7567)**, 47–55.

Becker, T., P. Bechtold, and I. Sandu, 2021: Characteristics of convective precipitation over tropical Africa in storm-resolving global simulations. *Quarterly Journal of the Royal Meteorological Society*, **147 (741)**, 4388–4407.

Bello, N., 1997: Investigating the spatial pattern of the characteristics of the onset and cessation of the rains in nigeria. *GeoJournal*, **43**, 113–123.

Benjamini, Y., and Y. Hochberg, 1995: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57 (1)**, 289–300.

Bi, K., L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian, 2022: Pangu-weather: A 3d high-resolution model for fast and accurate global weather forecast. *arXiv preprint arXiv:2211.02556*.

Bi, K., L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian, 2023: Accurate medium-range global weather forecasting with 3d neural networks. *Nature*, 1–6.

Bjerknes, V., 1904: Das problem der wettervorhersage, betrachtet vom standpunkte der mechanik und der physik. *Meteor. Z.*, **21**, 1–7.

Bolvin, D. T., G. J. Huffman, E. J. Nelkin, and J. Tan, 2021: Comparison of monthly imerg precipitation estimates with pacrain atoll observations. *Journal of Hydrometeorology*, **22 (7)**, 1745–1753.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Monthly weather review*, **78 (1)**, 1–3.

Burpee, R. W., 1972: The origin and structure of easterly waves in the lower troposphere of North Africa. *Journal of the Atmospheric Sciences*, **29 (1)**, 77–90.

Burpee, R. W., 1974: Characteristics of north african easterly waves during the summers of 1968 and 1969. *Journal of Atmospheric Sciences*, **31 (6)**, 1556–1570.

Burpee, R. W., 1975: Some features of synoptic–scale waves based on a compositing analysis of gate data. *Monthly Weather Review*, **103 (10)**, 921–925.

Cafaro, C., and Coauthors, 2021: Do convection-permitting ensembles lead to more skillful short-range probabilistic rainfall forecasts over tropical East Africa? *Weather and Forecasting*, **36 (2)**, 697–716.

Carlson, T. N., 1969: Synoptic histories of three african disturbances that developed into atlantic hurricanes. *Monthly Weather Review*, **97 (3)**, 256–276.

Chen, K., and Coauthors, 2023: FengWu: Pushing the skillful global medium-range weather forecast beyond 10 days lead. Preprint, https://arxiv.org/abs/2304.02948.

Chen, T., and C. Guestrin, 2016: Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.

Cheng, Y.-M., C. D. Thorncroft, and G. N. Kiladis, 2019: Two contrasting african easterly wave behaviors. *Journal of the Atmospheric Sciences*, **76 (6)**, 1753–1768.

Chollet, F., and Coauthors, 2015: Keras. https://keras.io.

Cook, K. H., and E. K. Vizy, 2019: Contemporary climate change of the african monsoon systems. *Current Climate Change Reports*, **5**, 145–159.

Davis, J., P. Knippertz, and A. H. Fink, 2013: The predictability of precipitation episodes during the West African dry season. *Quarterly Journal of the Royal Meteorological Society*, **139 (673)**, 1047–1058.

Dawid, A. P., 1984: Present position and potential developments: Some personal views statistical theory the prequential approach. *Journal of the Royal Statistical Society: Series A (General)*, **147 (2)**, 278–290.

de Andrade, F. M., M. P. Young, D. MacLeod, L. C. Hirons, S. J. Woolnough, and E. Black, 2021: Subseasonal precipitation prediction for Africa: Forecast evaluation and sources of predictability. *Weather and Forecasting*, **36 (1)**, 265–284.

Dias, J., M. Gehne, G. N. Kiladis, N. Sakaeda, P. Bechtold, and T. Haiden, 2018: Equatorial waves and the skill of ncep and ecmwf numerical weather prediction systems. *Monthly Weather Review*, **146 (6)**, 1763–1784.

Dickinson, M., and J. Molinari, 2000: Climatology of sign reversals of the meridional potential vorticity gradient over africa and australia. *Monthly weather review*, **128 (11)**, 3890–3900.

Diebold, F. X., and R. S. Mariano, 2002: Comparing predictive accuracy. *Journal of Business & Economic Statistics*, **20 (1)**, 134–144.

Diedhiou, A., S. Janicot, A. Viltard, P. De Felice, and H. Laurent, 1999: Easterly wave regimes and associated convection over west africa and tropical atlantic: Results from the ncep/ncar and ecmwf reanalyses. *Climate Dynamics*, **15**, 795–822.

Diez-Sierra, J., and M. del Jesus, 2020: Long-term rainfall prediction using atmospheric synoptic patterns in semi-arid climates with statistical and machine learning methods. *Journal of Hydrology*, **586**, 124 789.

Dimitriadis, T., T. Gneiting, and A. I. Jordan, 2021: Stable reliability diagrams for probabilistic classifiers. *Proceedings of the National Academy of Sciences*, **118 (8)**, e2016191 118.

Dodge, Y., 2003: *The Oxford dictionary of statistical terms*. OUP Oxford.

ECMWF Directorate, 2012: Describing ECMWF's forecasts and forecasting system. ECMWF Newsletter No. 133 – Autumn 2012. Available at http://www.ecmwf.int/sites/default/files/elibrary/2012/14576-newsletter-no133-autumn-2012.pdf.

Elless, T. J., and R. D. Torn, 2018: African easterly wave forecast verification and its relation to convective errors within the ECMWF ensemble prediction system. *Weather and Forecasting*, **33 (2)**, 461–477.

Feng, Z., and Coauthors, 2021: A global high-resolution mesoscale convective system database using satellite-derived cloud tops, surface precipitation, and tracking. *Journal of Geophysical Research: Atmospheres*, **126 (8)**, e2020JD034 202.

Fink, A., D. Vincent, and V. Ermert, 2006: Rainfall types in the West African Sudanian zone during the summer monsoon 2002. *Monthly weather review*, **134 (8)**, 2143–2164.

Fink, A. H., and A. Reiner, 2003: Spatiotemporal variability of the relation between African easterly waves and West African squall lines in 1998 and 1999. *Journal of Geophysical Research: Atmospheres*, **108 (D11)**, 1–17.

Fink, A. H., and Coauthors, 2011: Operational meteorology in West Africa: observational networks, weather analysis and forecasting. *Atmospheric Science Letters*, **12 (1)**, 135–141.

Fink, A. H., and Coauthors, 2017: Mean climate and seasonal cycle. *Meteorology of tropical West Africa: The forecasters' handbook*, 1–39.

Frank, N. L., 1969: The "inverted v" cloud pattern—an easterly wave? *Monthly Weather Review*, **97 (2)**, 130–140.

Friedman, J. H., 2001: Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189–1232.

Fröhlich, L., P. Knippertz, A. H. Fink, and E. Hohberger, 2013: An objective climatology of tropical plumes. *Journal of climate*, **26 (14)**, 5044–5060.

Gabriel, K., and J. Neumann, 1962: A Markov chain model for daily rainfall occurrence at Tel Aviv. *Quarterly Journal of the Royal Meteorological Society*, **88 (375)**, 90–95.

Garreaud, R. D., 2001: Subtropical cold surges: Regional aspects and global distribution. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, **21 (10)**, 1181–1197.

Gates, P., and H. Tong, 1976: On Markov chain modeling to some weather data. *Journal of Applied Meteorology and Climatology*, **15 (11)**, 1145–1151.

Gneiting, T., 2008: Probabilistic forecasting. Oxford University Press, 319–321 pp.

Gneiting, T., F. Balabdaoui, and A. E. Raftery, 2007: Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **69 (2)**, 243–268.

Gneiting, T., and M. Katzfuss, 2014: Probabilistic forecasting. *Annual Review of Statistics and Its Application*, **1**, 125–151.

Gneiting, T., and A. E. Raftery, 2007: Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, **102 (477)**, 359–378.

Gneiting, T., A. E. Raftery, A. H. Westveld, and T. Goldman, 2005: Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. *Monthly Weather Review*, **133 (5)**, 1098–1118.

Gneiting, T., and E.-M. Walz, 2022: Receiver operating characteristic (ROC) movies, universal ROC (UROC) curves, and coefficient of predictive ability (CPA). *Machine Learning*, **111 (8)**, 2769–2797.

Gu, G., 2009: Intraseasonal variability in the equatorial atlantic-west africa during march–june. *Climate dynamics*, **32**, 457–471.

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez, 2009: Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. *Journal of hydrology*, **377 (1-2)**, 80–91.

Haiden, T., M. J. Rodwell, D. S. Richardson, A. Okagaki, T. Robinson, and T. Hewson, 2012: Intercomparison of global model precipitation forecast skill in 2010/11 using the SEEPS score. *Monthly Weather Review*, **140 (8)**, 2720–2733.

Hamill, T. M., and S. J. Colucci, 1997: Verification of eta–rsm short-range ensemble forecasts. *Monthly Weather Review*, **125 (6)**, 1312–1327.

Hamilton, R., J. Archbold, and C. Douglas, 1945: Meteorology of nigeria and adjacent territory. *Quarterly Journal of the Royal Meteorological Society*, **71 (309-310)**, 231–264.

Hanley, K. E., J. S. Pirret, C. L. Bain, A. J. Hartley, H. W. Lean, S. Webster, and B. J. Woodhams, 2021: Assessment of convection-permitting versions of the unified model over the Lake Victoria basin region. *Quarterly Journal of the Royal Meteorological Society*, **147 (736)**, 1642–1660.

Hendon, H. H., and K. Woodberry, 1993: The diurnal cycle of tropical convection. *Journal of Geophysical Research: Atmospheres*, **98 (D9)**, 16 623–16 637.

Henzi, A., J. F. Ziegel, and T. Gneiting, 2021: Isotonic distributional regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **83**, 963–993.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146 (730)**, 1999–2049.

Hewson, T. D., and F. M. Pillosu, 2021: A low-cost post-processing technique improves weather forecasts around the world. *Communications Earth & Environment*, **2 (1)**, 132.

Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement mission. *Bulletin of the American Meteorological Society*, **97 (5)**, 701–722.

Huffman, G. J., and Coauthors, 2020: NASA Global Precipitation Measurement (GPM) Integrated Mult-satellitE Retrievals for GPM (IMERG), Algorithm Theoretical Basis Document (ATBD) Version 06. NASA, URL https://gpm.nasa.gov/sites/default/files/2020-05/IMERG_ATBD_V06.3.pdf.

Janicot, S., F. Mounier, S. Gervois, B. Sultan, and G. N. Kiladis, 2010: The dynamics of the west african monsoon. part v: The detection and role of the dominant modes of convectively coupled equatorial rossby waves. *Journal of Climate*, **23 (14)**, 4005–4024.

Janiga, M. A., C. J. Schreck, J. A. Ridout, M. Flatau, N. P. Barton, E. J. Metzger, and C. A. Reynolds, 2018: Subseasonal forecasts of convectively coupled equatorial waves and the mjo: Activity and predictive skill. *Monthly Weather Review*, **146 (8)**, 2337–2360.

Jørgensen, B., 1987: Exponential dispersion models. *Journal of the Royal Statistical Society: Series B (Methodological)*, **49 (2)**, 127–145.

Judt, F., 2020: Atmospheric predictability of the tropics, middle latitudes, and polar regions explored through global storm-resolving simulations. *Journal of the Atmospheric Sciences*, **77 (1)**, 257–276.

Jung, H., and P. Knippertz, 2023: Link between the time-space behavior of rainfall and 3d dynamical structures of equatorial waves in global convection-permitting simulations. *Geophysical Research Letters*, **50 (2)**, e2022GL100 973.

Kiladis, G. N., C. D. Thorncroft, and N. M. Hall, 2006: Three-dimensional structure and dynamics of African easterly waves. part i: Observations. *Journal of Atmospheric Sciences*, **63 (9)**, 2212–2230.

Kiladis, G. N., and K. M. Weickmann, 1997: Horizontal structure and seasonality of large-scale circulations associated with submonthly tropical convection. *Monthly weather review*, **125 (9)**, 1997–2013.

Kiladis, G. N., M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy, 2009: Convectively coupled equatorial waves. *Reviews of Geophysics*, **47 (2)**.

Kingma, D. P., and J. Ba, 2014: Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kniffka, A., and Coauthors, 2020: An evaluation of operational and research weather forecasts for southern West Africa using observations from the DACCIWA field campaign in June–July 2016. *Quarterly Journal of the Royal Meteorological Society*, **146 (728)**, 1121–1148.

Knippertz, P., 2003: Tropical-extratropical interactions causing precipitation in northwest Africa: Statistical analysis and seasonal variations. *Monthly Weather Review*, **131 (12)**, 3069–3076.

Knippertz, P., 2007: Tropical–extratropical interactions related to upper-level troughs at low latitudes. *Dynamics of Atmospheres and Oceans*, **43 (1-2)**, 36–62.

Knippertz, P., and A. H. Fink, 2009: Prediction of dry-season precipitation in tropical West Africa and its relation to forcing from the extratropics. *Weather and Forecasting*, **24 (4)**, 1064–1084.

Knippertz, P., A. H. Fink, A. Reiner, and P. Speth, 2003: Three late summer/early autumn cases of tropical-extratropical interactions causing precipitation in Northwest Africa. *Monthly Weather Review*, **131 (1)**, 116–135.

Knippertz, P., and J. E. Martin, 2005: Tropical plumes and extreme precipitation in subtropical and tropical West Africa. *Quarterly Journal of the Royal Meteorological Society*, **131 (610)**, 2337–2365.

Knippertz, P., and Coauthors, 2017: A meteorological and chemical overview of the DACCIWA field campaign in West Africa in June–July 2016. *Atmospheric Chemistry and Physics*, **17 (17)**, 10 893–10 918.

Knippertz, P., and Coauthors, 2022: The intricacies of identifying equatorial waves. *Quarterly Journal of the Royal Meteorological Society*, **148 (747)**, 2814–2852.

Knoben, W. J., J. E. Freer, and R. A. Woods, 2019: Inherent benchmark or not? comparing nash–sutcliffe and kling–gupta efficiency scores. *Hydrology and Earth System Sciences*, **23 (10)**, 4323–4331.

Lafore, J. P., and Coauthors, 2017: Deep convection. *Meteorology of Tropical West Africa: The Forecasters' Handbook*, D. J. Parker, and M. Diop-Kane, Eds., Wiley, Chichester, 90–129.

Laing, A. G., R. E. Carbone, and V. Levizzani, 2011: Cycles and propagation of deep convection over equatorial africa. *Monthly Weather Review*, **139 (9)**, 2832–2853.

Lam, R., and Coauthors, 2022: GraphCast: Learning skillful medium-range global weather forecasting. Preprint, https://arxiv.org/abs/2212.12794.

Lavaysse, C., A. Diedhiou, H. Laurent, and T. Lebel, 2006: African easterly waves and convective activity in wet and dry sequences of the West African Monsoon. *Climate Dynamics*, **27 (2)**, 319–332.

Lavender, S. L., and A. J. Matthews, 2009: Response of the west african monsoon to the madden–julian oscillation. *Journal of Climate*, **22 (15)**, 4097–4116.

Lawton, Q. A., S. J. Majumdar, K. Dotterer, C. Thorncroft, and C. J. Schreck III, 2022: The influence of convectively coupled kelvin waves on african easterly waves in a wave-following framework. *Monthly Weather Review*, **150 (8)**, 2055–2072.

Lazo, J. K., R. E. Morss, and J. L. Demuth, 2009: 300 billion served: Sources, perceptions, uses, and values of weather forecasts. *Bulletin of the American Meteorological Society*, **90 (6)**, 785–798.

Li, Y., and S. N. Stechmann, 2020: Predictability of tropical rainfall and waves: Estimates from observational data. *Quarterly Journal of the Royal Meteorological Society*, **146 (729)**, 1668–1684.

Liu, Y., and Coauthors, 2016: Application of deep convolutional neural networks for detecting extreme weather in climate datasets. *arXiv preprint arXiv:1605.01156*.

Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical pacific. *Journal of Atmospheric Sciences*, **28 (5)**, 702–708.

Madden, R. A., and P. R. Julian, 1972: Description of global-scale circulation cells in the tropics with a 40–50 day period. *Journal of Atmospheric Sciences*, **29 (6)**, 1109–1123.

Maier-Gerber, M., 2023: Subseasonal tropical cyclone activity over the north atlantic: A systematic comparison of different forecast approaches.

Maier-Gerber, M., A. H. Fink, M. Riemer, E. Schoemer, C. Fischer, and B. Schulz, 2021: Statistical–dynamical forecasting of subseasonal North Atlantic tropical cyclone occurrence. *Weather and Forecasting*, **36 (6)**, 2127–2142.

Maranan, M., A. H. Fink, and P. Knippertz, 2018: Rainfall types over southern West Africa: Objective identification, climatology and synoptic environment. *Quarterly Journal of the Royal Meteorological Society*, **144 (714)**, 1628–1648.

Marsham, J. H., P. Knippertz, N. S. Dixon, D. J. Parker, and G. M. Lister, 2011: The importance of the representation of deep convection for modeled dust-generating winds over West Africa during summer. *Geophysical Research Letters*, **38 (16)**, L16 803.

Matheson, J. E., and R. L. Winkler, 1976: Scoring rules for continuous probability distributions. *Management science*, **22 (10)**, 1087–1096.

Mathon, V., H. Laurent, and T. Lebel, 2002: Mesoscale convective system rainfall in the Sahel. *Journal of Applied Meteorology and Climatology*, **41 (11)**, 1081–1092.

Matsuno, T., 1966: Quasi-geostrophic motions in the equatorial area. *Journal of the Meteorological Society of Japan. Ser. II*, **44 (1)**, 25–43.

Matthews, A. J., 2000: Propagation mechanisms for the madden-julian oscillation. *Quarterly Journal of the Royal Meteorological Society*, **126 (569)**, 2637–2651.

Matthews, A. J., 2004: Intraseasonal variability over tropical africa during northern summer. *Journal of Climate*, **17 (12)**, 2427–2440.

McCulloch, C. E., 2000: Generalized linear models. *Journal of the American Statistical Association*, **95 (452)**, 1320–1324.

Mcguirk, J. P., A. H. Thompson, and N. R. Smith, 1987: Moisture bursts over the tropical pacific ocean. *Monthly weather review*, **115 (4)**, 787–798.

McGuirk, J. P., and D. J. Ulsh, 1990: Evolution of tropical plumes in vas water vapor imagery. *Monthly weather review*, **118 (9)**, 1758–1766.

Mekonnen, A., C. J. Schreck III, and B. D. Enyew, 2020: The impact of kelvin wave activity during dry and wet african summer rainfall years. *Atmosphere*, **11 (6)**, 568.

Mekonnen, A., and C. D. Thorncroft, 2016: On mechanisms that determine synoptic time scale convection over east africa. *International Journal of Climatology*, **36 (12)**, 4045–4057.

Mekonnen, A., C. D. Thorncroft, and A. R. Aiyyer, 2006: Analysis of convection and its association with African easterly waves. *Journal of Climate*, **19 (20)**, 5405–5421.

Mounier, F., G. N. Kiladis, and S. Janicot, 2007: Analysis of the dominant mode of convectively coupled kelvin waves in the west african monsoon. *Journal of Climate*, **20 (8)**, 1487–1503.

Murphy, A. H., 1977: The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation. *Monthly Weather Review*, **105 (7)**, 803–816.

Nelder, J. A., and R. W. Wedderburn, 1972: Generalized linear models. *Journal of the Royal Statistical Society Series A: Statistics in Society*, **135 (3)**, 370–384.

Nguyen, H., and J.-P. Duvel, 2008: Synoptic wave perturbations and convective systems over equatorial africa. *Journal of Climate*, **21 (23)**, 6372–6388.

Nguyen, T., J. Brandstetter, A. Kapoor, J. K. Gupta, and A. Grover, 2023: ClimaX: A foundation model for weather and climate. Preprint, https://arxiv.org/abs/2301.10343.

Nicholson, S. E., 2008: The intensity, location and structure of the tropical rainbelt over West Africa as factors in interannual variability. *International Journal of Climatology*, **28 (13)**, 1775–1785.

Nicholson, S. E., 2009: On the factors modulating the intensity of the tropical rainbelt over West Africa. *International Journal of Climatology*, **29 (5)**, 673–689.

Nielsen, M. A., 2015: *Neural networks and deep learning*, Vol. 25. Determination press San Francisco, CA, USA.

Palmer, T. N., 2002: The economic value of ensemble forecasts as a tool for risk assessment: From days to decades. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, **128 (581)**, 747–774.

Pante, G., and P. Knippertz, 2019: Resolving Sahelian thunderstorms improves mid-latitude weather forecasts. *Nature Communications*, **10**, 3487.

Parker, D. J., and Coauthors, 2008: The amma radiosonde program and its implications for the future of atmospheric monitoring over africa. *Bulletin of the American Meteorological Society*, **89 (7)**, 1015–1028.

Pathak, J., and Coauthors, 2022: Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. *arXiv preprint arXiv:2202.11214*.

Pedregosa, F., and Coauthors, 2011a: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12 (Oct)**, 2825–2830.

Pedregosa, F., and Coauthors, 2011b: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830.

Peyrille, P., J.-P. Lafore, and A. Boone, 2016: The annual cycle of the west african monsoon in a two-dimensional model: Mechanisms of the rain-band migration. *Quarterly Journal of the Royal Meteorological Society*, **142 (696)**, 1473–1489.

Peyrillé, P., R. Roehrig, and S. Sanogo, 2023: Tropical waves are key drivers of extreme precipitation events in the central sahel. *Geophysical Research Letters*, **50 (20)**, e2023GL103 715.

Piersig, W., 1936: *Schwankungen von Luftdruck und Luftbewegung: sowie ein Beitrag zum Wettergeschehen im Passatgebiet des östlichen Nordatlantischen Ozeans*. Deutsche Seewarte [publisher].

Pohl, B., S. Janicot, B. Fontaine, and R. Marteau, 2009: Implication of the madden–julian oscillation in the 40-day variability of the west african monsoon. *Journal of Climate*, **22 (13)**, 3769–3785.

Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski, 2005: Using bayesian model averaging to calibrate forecast ensembles. *Monthly weather review*, **133 (5)**, 1155–1174.

Raj, J., H. K. Bangalath, and G. Stenchikov, 2019: West african monsoon: current state and future projections in a high-resolution agcm. *Climate dynamics*, **52 (11)**, 6441–6461.

Rasheeda Satheesh, A., P. Knippertz, A. H. Fink, E.-M. Walz, and T. Gneiting, 2023: Sources of predictability of synoptic-scale rainfall during the west african summer monsoon. *Quarterly Journal of the Royal Meteorological Society*.

Ravuri, S., and Coauthors, 2021: Skilful precipitation nowcasting using deep generative models of radar. *Nature*, **597 (7878)**, 672–677.

Reed, R. J., D. C. Norquist, and E. E. Recker, 1977: The structure and properties of African wave disturbances as observed during phase III of GATE. *Monthly Weather Review*, **105 (3)**, 317–333.

Regula, H., 1936: *Druckschwankungen und Tornados an der Westküste von Afrika.*

Ribeiro, M. T., S. Singh, and C. Guestrin, 2016: " why should i trust you?" explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 1135–1144.

Roehrig, R., D. Bouniol, F. Guichard, F. Hourdin, and J.-L. Redelsperger, 2013: The present and future of the west african monsoon: A process-oriented assessment of cmip5 simulations along the amma transect. *Journal of Climate*, **26 (17)**, 6471–6505.

Roehrig, R., F. Chauvin, and J.-P. Lafore, 2011: 10–25-day intraseasonal variability of convection over the sahel: A role of the saharan heat low and midlatitudes. *Journal of climate*, **24 (22)**, 5863–5878.

Roundy, P. E., 2012: Tracking and prediction of large-scale organized tropical convection by spectrally focused two-step space–time eof analysis. *Quarterly Journal of the Royal Meteorological Society*, **138 (665)**, 919–931.

Roundy, P. E., C. J. Schreck, and M. A. Janiga, 2009: Contributions of convectively coupled equatorial rossby waves and kelvin waves to the real-time multivariate mjo indices. *Monthly Weather Review*, **137 (1)**, 469–478.

Sanders, F., 1963: On subjective probability forecasting. *Journal of Applied Meteorology and Climatology*, **2 (2)**, 191–201.

Scheuerer, M., 2014: Probabilistic quantitative precipitation forecasting using ensemble model output statistics. *Quarterly Journal of the Royal Meteorological Society*, **140 (680)**, 1086–1096.

Scheuerer, M., and T. M. Hamill, 2015: Statistical postprocessing of ensemble precipitation forecasts by fitting censored, shifted gamma distributions. *Monthly Weather Review*, **143 (11)**, 4578–4596.

Schlueter, A., A. H. Fink, and P. Knippertz, 2019a: A systematic comparison of tropical waves over northern Africa. Part II: Dynamics and thermodynamics. *Journal of Climate*, **32 (9)**, 2605–2625.

Schlueter, A., A. H. Fink, P. Knippertz, and P. Vogel, 2019b: A systematic comparison of tropical waves over northern Africa. Part I: Influence on rainfall. *Journal of Climate*, **32 (5)**, 1501–1523.

Schlüter, A., 2019: Tropical waves and rainfall over africa: Variability, mechanisms and potential for forecasting.

Schulz, B., 2023: Statistical postprocessing of numerical weather prediction forecasts using machine learning.

Senior, C. A., and Coauthors, 2021: Convection-permitting regional climate change simulations for understanding future climate and informing decision-making in Africa. *Bulletin of the American Meteorological Society*, **102 (6)**, E1206–E1223.

Shuman, F. G., 1989: History of numerical weather prediction at the national meteorological center. *Weather and forecasting*, **4 (3)**, 286–296.

Sinclaire, Z., A. Lenouo, C. Tchawoua, and S. Janicot, 2015: Synoptic kelvin type perturbation waves over congo basin over the period 1979–2010. *Journal of Atmospheric and Solar-Terrestrial Physics*, **130**, 43–56.

Sloughter, J. M. L., A. E. Raftery, T. Gneiting, and C. Fraley, 2007: Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *Monthly Weather Review*, **135 (9)**, 3209–3220.

Söhne, N., J.-P. Chaboureau, and F. Guichard, 2008: Verification of cloud cover forecast with satellite observation over west africa. *Monthly Weather Review*, **136 (11)**, 4421–4434.

Štrumbelj, E., and I. Kononenko, 2014: Explaining prediction models and individual predictions with feature contributions. *Knowledge and information systems*, **41**, 647–665.

Sultan, B., and S. Janicot, 2000: Abrupt shift of the itcz over west africa and intraseasonal variability. *Geophysical Research Letters*, **27 (20)**, 3353–3356.

Sultan, B., and S. Janicot, 2003: The west african monsoon dynamics. part ii: The "preonset" and "onset" of the summer monsoon. *Journal of climate*, **16 (21)**, 3407–3427.

Takayabu, Y. N., and T. Nitta, 1993: 3-5 day-period disturbances coupled with convection over the tropical pacific ocean. *Journal of the Meteorological Society of Japan. Ser. II*, **71 (2)**, 221–246.

Taylor, C. M., R. A. de Jeu, F. Guichard, P. P. Harris, and W. A. Dorigo, 2012: Afternoon rain more likely over drier soils. *Nature*, **489 (7416)**, 423–426.

Taylor, C. M., and Coauthors, 2022: Nowcasting tracks of severe convective storms in west africa from observations of land surface state. *Environmental Research Letters*, **17 (3)**, 034 016.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of geophysical research: atmospheres*, **106 (D7)**, 7183–7192.

Thorncroft, C., and B. Hoskins, 1994: An idealized study of African easterly waves. i: A linear view. *Quarterly Journal of the Royal Meteorological Society*, **120 (518)**, 953–982.

Tindall, J., J. Thuburn, and E. Highwood, 2006: Equatorial waves in the lower stratosphere. i: A novel detection method. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, **132 (614)**, 177–194.

Tsantekidis, A., N. Passalis, A. Tefas, J. Kanniainen, M. Gabbouj, and A. Iosifidis, 2017: Forecasting stock prices from the limit order book using convolutional neural networks. *2017 IEEE 19th conference on business informatics (CBI)*, IEEE, Vol. 1, 7–12.

Tulich, S. N., and G. N. Kiladis, 2012: Squall lines and convectively coupled gravity waves in the tropics: Why do most cloud systems propagate westward? *Journal of the Atmospheric Sciences*, **69 (10)**, 2995–3012.

United Nations, Department of Economic and Social Affairs, Population Division, 2022: World population prospects 2022: Methodology of the united nations population estimates and projections. Technical Report UN DESA/POP/2022/TR/NO. 4, UN DESA.

van der Linden, R., A. H. Fink, J. G. Pinto, T. Phan-Van, and G. N. Kiladis, 2016: Modulation of daily rainfall in southern Vietnam by the Madden–Julian oscillation and convectively coupled equatorial waves. *Journal of Climate*, **29 (16)**, 5801–5820.

van der Linden, R., P. Knippertz, A. H. Fink, B. Ingleby, M. Maranan, and A. Benedetti, 2020: The influence of DACCIWA radiosonde data on the quality of ECMWF analyses and forecasts over southern West Africa. *Quarterly Journal of the Royal Meteorological Society*, **146 (729)**, 1719–1739.

Vannitsem, S., D. S. Wilks, and J. Messner, Eds., 2018: *Statistical Postprocessing of Ensemble Forecasts*. Elsevier, Amsterdam.

Ventrice, M. J., and C. D. Thorncroft, 2013: The role of convectively coupled atmospheric Kelvin waves on African easterly wave activity. *Monthly weather review*, **141 (6)**, 1910–1924.

Vizy, E. K., and K. H. Cook, 2009: A mechanism for african monsoon breaks: Mediterranean cold air surges. *Journal of Geophysical Research: Atmospheres*, **114 (D1)**.

Vizy, E. K., and K. H. Cook, 2014: Impact of cold air surges on rainfall variability in the sahel and wet african tropics: A multi-scale analysis. *Climate dynamics*, **43**, 1057–1081.

Vogel, P., P. Knippertz, A. H. Fink, A. Schlueter, and T. Gneiting, 2018: Skill of global raw and postprocessed ensemble predictions of rainfall over northern tropical Africa. *Weather and Forecasting*, **33 (2)**, 369–388.

Vogel, P., P. Knippertz, A. H. Fink, A. Schlueter, and T. Gneiting, 2020: Skill of global raw and postprocessed ensemble predictions of rainfall in the tropics. *Weather and Forecasting*, **35 (6)**, 2367–2385.

Vogel, P., P. Knippertz, T. Gneiting, A. H. Fink, M. Klar, and A. Schlueter, 2021: Statistical forecasts for the occurrence of precipitation outperform global models over northern tropical Africa. *Geophysical Research Letters*, **48 (3)**, e2020GL091 022.

Vollmert, P., A. H. Fink, and H. Besler, 2003: "Ghana dry zone" und "Dahomey gap": Ursachen für eine niederschlagsanomalie im tropischen westafrika. *Erde*, **134 (4)**, 375–393.

Walz, E.-M., A. Henzi, J. Ziegel, and T. Gneiting, 2022: Easy uncertainty quantification (easyuq): Generating predictive distributions from single-valued model output. *arXiv preprint arXiv:2212.08376*.

Wang, H., and R. Fu, 2007: The influence of amazon rainfall on the atlantic itcz through convectively coupled kelvin waves. *Journal of Climate*, **20 (7)**, 1188–1201.

Wang, Y., and Coauthors, 2023: Verification of operational numerical weather prediction model forecasts of precipitation using satellite rainfall estimates over Africa. *Meteorological Applications*, **30**, e2112.

Warner, J. L., J. Petch, C. J. Short, and C. Bain, 2023: Assessing the impact of a NWP warm-start system on model spin-up over tropical Africa. *Quarterly Journal of the Royal Meteorological Society*, **149 (751)**, 621–636.

Wheeler, M., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber–frequency domain. *Journal of Atmospheric Sciences*, **56 (3)**, 374–399.

Wheeler, M., and K. M. Weickmann, 2001: Real-time monitoring and prediction of modes of coherent synoptic to intraseasonal tropical variability. *Monthly Weather Review*, **129 (11)**, 2677–2694.

Wilks, D. S., 1990: On the combination of forecast probabilities for consecutive precipitation periods. *Weather and forecasting*, **5 (4)**, 640–650.

Wilks, D. S., 2011: *Statistical methods in the atmospheric sciences*, Vol. 100. Academic press.

Wilks, D. S., 2016: "The stippling shows statistically significant grid points": How research results are routinely overstated and overinterpreted, and what to do about it. *Bulletin of the American Meteorological Society*, **97 (12)**, 2263–2273.

Woodhams, B. J., C. E. Birch, J. H. Marsham, C. L. Bain, N. M. Roberts, and D. F. Boyd, 2018: What is the added value of a convection-permitting model for

forecasting extreme rainfall over tropical East Africa? *Monthly Weather Review*, **146 (9)**, 2757–2780.

Yang, G.-Y., B. Hoskins, and J. Slingo, 2003: Convectively coupled equatorial waves: A new methodology for identifying wave structures in observational data. *Journal of the atmospheric sciences*, **60 (14)**, 1637–1654.

Yang, G.-Y., J. Methven, S. Woolnough, K. Hodges, and B. Hoskins, 2018: Linking african easterly wave activity with equatorial waves and the influence of rossby waves from the southern hemisphere. *Journal of the Atmospheric Sciences*, **75 (6)**, 1783–1809.

Yasunaga, K., and B. Mapes, 2012: Differences between more divergent and more rotational types of convectively coupled equatorial waves. part II: Composite analysis based on space–time filtering. *Journal of the atmospheric sciences*, **69 (1)**, 17–34.

Zadrozny, B., and C. Elkan, 2002: Transforming classifier scores into accurate multiclass probability estimates. *KDD '02: Proceedings of the eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 694–699.

Zamo, M., and P. Naveau, 2018: Estimation of the continuous ranked probability score with limited information and applications to ensemble weather forecasts. *Mathematical Geosciences*, **50 (2)**, 209–234.

Zhang, C., 2005: Madden-julian oscillation. *Reviews of Geophysics*, **43 (2)**.

Zhu, C., R. H. Byrd, P. Lu, and J. Nocedal, 1997: Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on mathematical software (TOMS)*, **23 (4)**, 550–560.

# Appendices

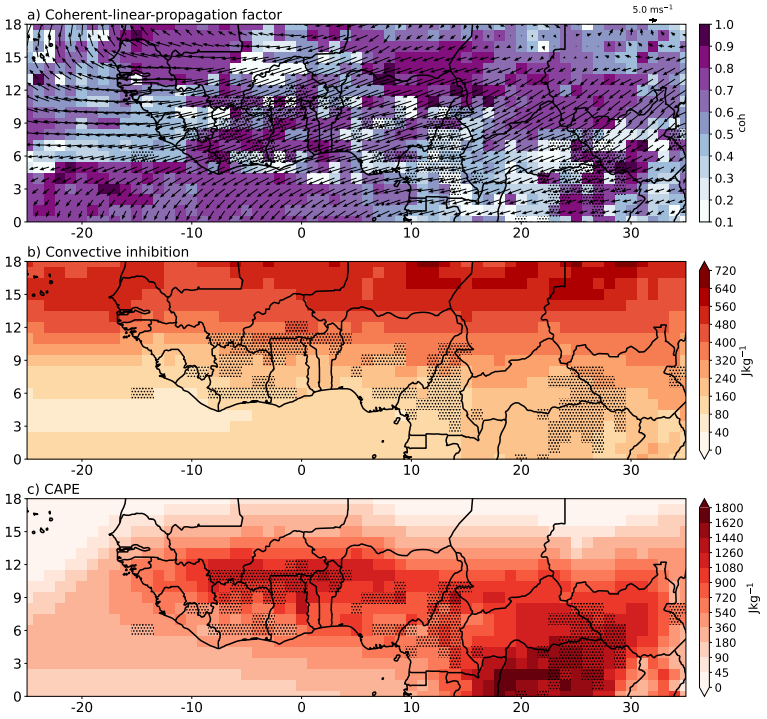



**Figure A1:** Same as Fig. 6.2 but for the MJ season.

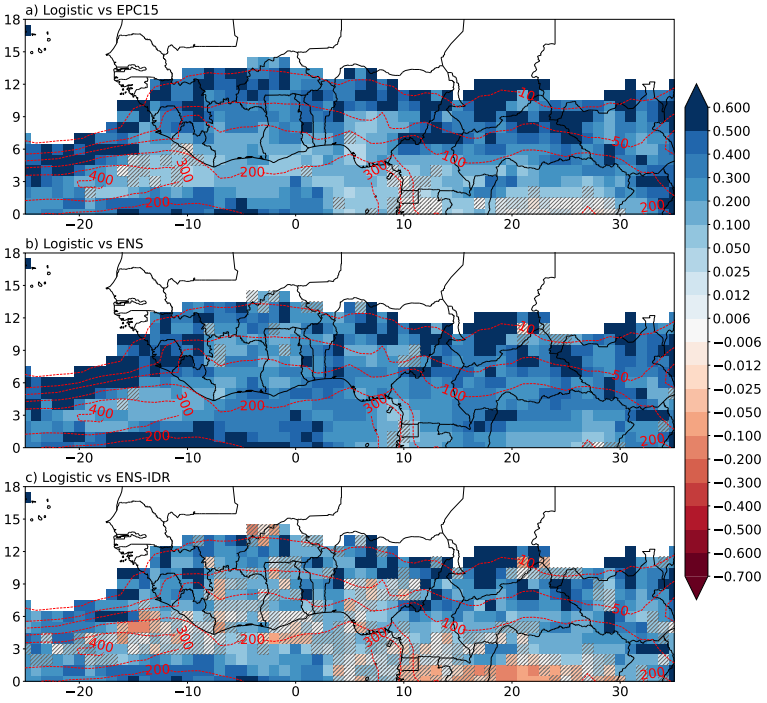



**Figure A2:** Same as Fig. 6.7 but for the MA season.

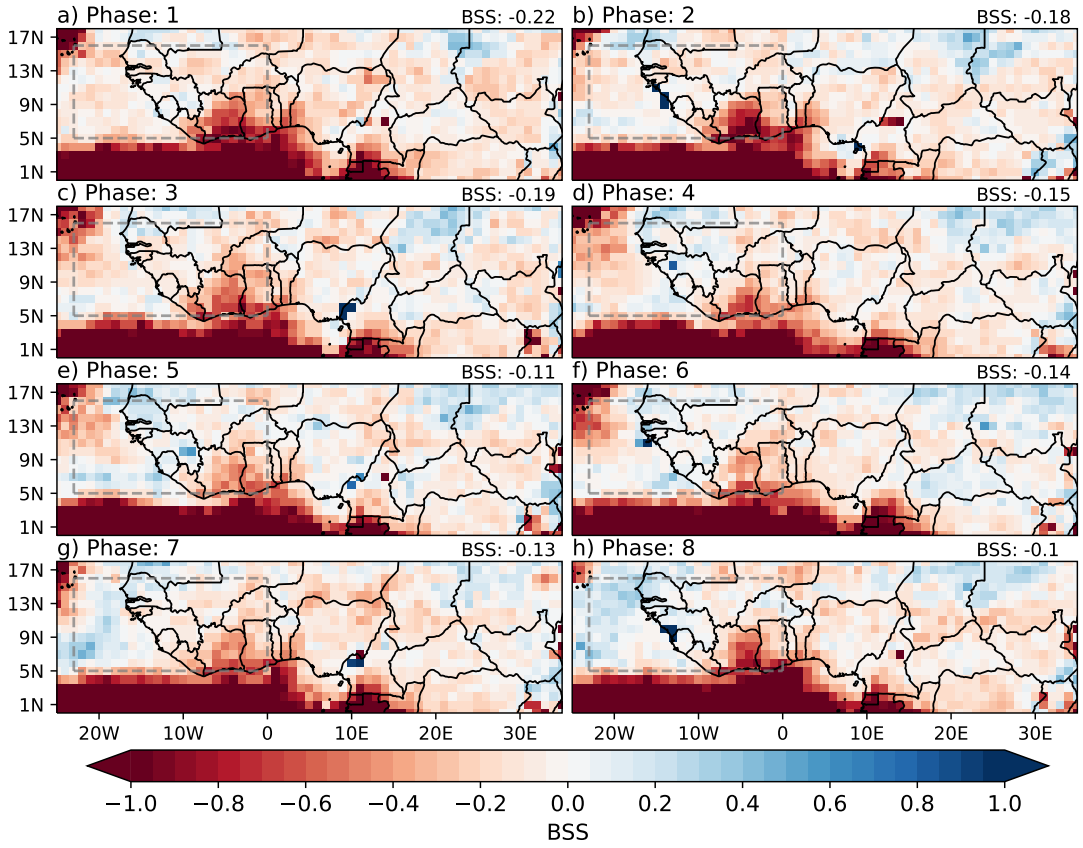



**Figure A3:** Same as Fig. 7.13 but for MRG.

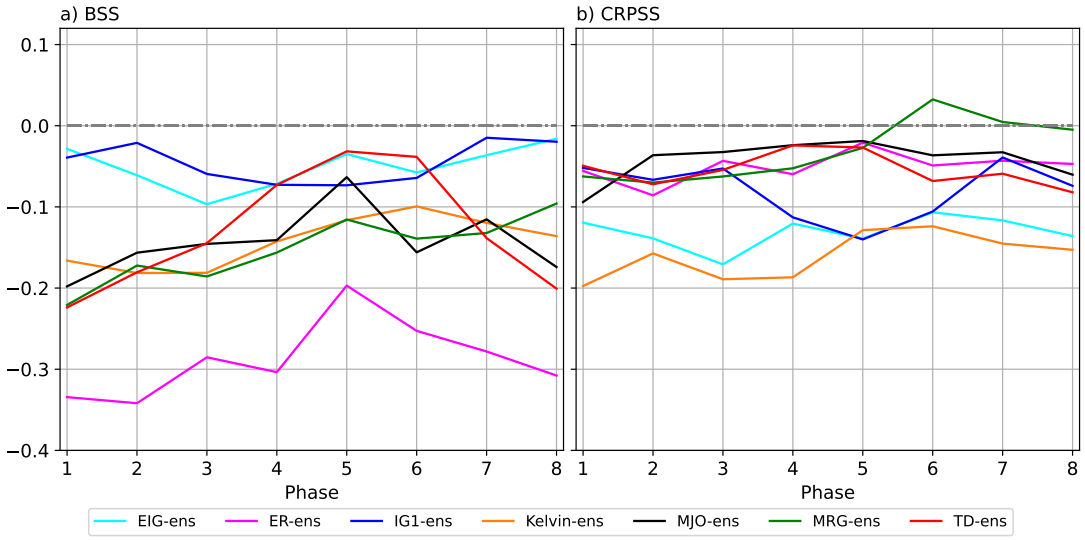



**Figure A4:** Same as Fig. 7.12 but for an amplitude threshold of 0.5 standard deviation

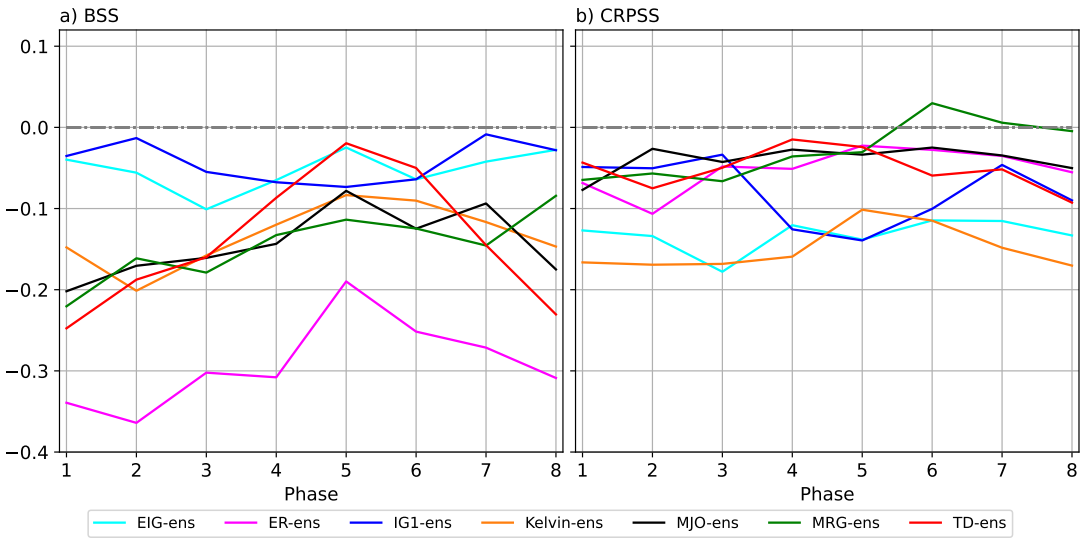



**Figure A5:** Same as Fig. 7.12 but for an amplitude threshold of 1.0 standard deviation

# Acknowledgements

I want to express my heartfelt gratitude to my advisors, Peter Knippertz and Andreas H. Fink, whose unwavering support and encouragement have been the bedrock of my entire PhD journey. Your invaluable feedback and constructive criticism have played a pivotal role in shaping and refining my work, leading to the successful completion of this thesis. Our Friday discussions have always been a source of intellectual stimulation. These discussions have not only enhanced my research skills but have also elevated my own standards, setting a benchmark for whatever path my career may set course for in the future. I am truly grateful for the mentorship that goes beyond professional guidance, making these past four years not just academically enriching but also personally fulfilling.

I extend my heartfelt gratitude to Tilmann Gneiting, my third advisor, whose unwavering support and insightful suggestions have been invaluable to my project. A special thanks goes to my project partner, Eva, whose assistance with essential tools and software has been the cornerstone of my project's success, and I am truly grateful for our partnership. I am indebted to Waves-to-Weather for creating a shared space under their umbrella. The events and workshops they organized not only expanded our skill sets but also provided a platform to connect with extraordinary scientists, both seasoned and emerging. I also want to express my appreciation to my ex-colleagues Roderick and Gregor. Their initial assistance was instrumental in getting my PhD off the ground, and I'm deeply grateful for their support.

I'm incredibly grateful to Marlon for crafting the most seamless first day at work for me. From warm introductions to navigating through the intimidating paperwork, his helping hand made the onboarding process a breeze. I thank you for being the go-to person for anything socialising-related and for sharing a common passion for bouldering. In essence, Marlon has not only been a great colleague but also a friend who has made my work experience truly enjoyable.

I feel incredibly fortunate to have shared my office space with two amazing individuals, Christoph and Raffi. Christoph, your enthusiasm for diverse discussions has added a unique flavour to our workdays. I'm grateful for your unwavering support during code-debugging sessions, and for always being ready to tackle challenges together. Raffi, your open-mindedness and warmth have made our workspace a truly welcoming environment. Thank you for introducing me to the world of bouldering, a passion that I'm sure will

stick with me for the rest of my life. I thank Beth for being the kindest person on the floor, for always being ready to help, and especially for baking cakes for us from time to time. I thank Lea for being the 'organizer' on the floor. But more than that, I thank her for being a great friend who was more excited than me for my first Mario Kart experience. I feel immensely grateful to be able to call all of you my friends.

The journey through my PhD would have been unimaginably different without the presence of Hyunju and Maurus. Reflecting on the countless memories we've shared together, it's the moments of extreme confusion that stand out the most, a testament to the depth of our bond and the challenges we navigated. Now, when I think about those times, I realize that I wouldn't trade even a single second of those very confusing times for anything else in the world. Hyunju and Maurus, your friendship has been a guiding light, turning the 'trauma' of academia into shared laughter. Thank you for being not just colleagues but two of my dearest friends.

I want to express my heartfelt gratitude to Hyunju, Marlon, Beth, Gabi, and Chetan for reviewing parts of my thesis. Your insights and suggestions have been truly invaluable.

I extend my heartfelt gratitude to my family for their regular check-in calls over the past four years, especially during the pandemic, even though we've been separated by three different time zones. A special thank you goes to my mother, whose tireless efforts in providing for me until I began earning, along with her continuous encouragement to dream big, have been the inspiration behind my journey. Her ability to always find ways to support my aspirations, despite being a single mother, has played a pivotal role in shaping the person I am today. I am profoundly grateful for her relentless struggle and the indelible mark it has left on my life.

Finally, I'm immensely lucky to have Gabi in my life. Thank you for constantly being there for me providing unwavering support during my low moments, especially during thesis writing. Thank you for willingly lending an ear to my rants. But above all this, thank you from the bottom of my heart for being an extraordinary partner.