MODELING THE SPACE-TIME VARIABILITY OF HYDROCLIMATE OVER THE MONSOON REGIONS OF EAST AND WEST AFRICA FOR WATER RESOURCES AND LIVESTOCK MANAGEMENT

by

DANIEL PAUL BROMAN

B.S., University of New Hampshire, 2011

M.S., University of Colorado, 2013

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Civil Engineering

2019

This thesis entitled:
Modeling the Space-Time Variability of Hydroclimate over the Monsoon Regions of East and West Africa for Water Resources and Livestock Management
written by Daniel Paul Broman
has been approved for the Department of Civil Engineering

Dr. Balaji Rajagopalan (Committee Chair)

Dr. Thomas Hopson (Committee Member)

Dr. Edith Zagona (Committee Member)

Dr. Ben Livneh (Committee Member)

Dr. Jean-Francois Lamarque (Committee Member)

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Broman, Daniel Paul (Ph.D., Civil Engineering)
Modeling the Space-Time Variability of Hydroclimate over the Monsoon Regions of East and West Africa for Water Resources and Livestock Management
Thesis directed by Professor Balaji Rajagopalan

There has been an increasing interest in and need for the application of climate information to address societal issues. Partially arising from the increasing confidence in anthropogenic climate change (Stocker et al. 2013) and partially from increasingly complex and resource-limited conditions, climate impacts are now considered in fields as diverse as natural hazards, public health, and agriculture. This research looks to show the utility of existing data analysis and statistical tools in providing actionable information to decision makers. Though this information is relevant in a wide variety of locations and conditions, the focus of this research has been on producing information for applications in the developing world. Climate variability at multiple timescales can have an outsized impact in the developing world given the often limited infrastructure, and weaker social and economic institutions (Field et al. 2014). There is a need for tools and analyses that translate climate information at coarse space and time scales to local scales where decisions of resource management are made. Through climate diagnostics, precipitation associated with the West African and East African monsoons show variations both subseasonally and interseasonally over the 20th and early 21st centuries. While prominent events including a wet period in the middle of the 20th century and a drought in the late 20th century have been well documented, this research highlights how this interseasonal variability can be linked to changes in large-scale teleconnections. In examining these teleconnections this research also notes that these teleconnections have entered into a new epoch. Tools that translate this information to address local issues have been developed. In West Africa, a stochastic weather generator can simulate the weather variables required to assess livestock heat stress and offer a means of forecasting and spatial modeling. This information can provide pastoralists

better information during their seasonal migration. In East Africa, an understanding of precipitation variability and its coupling with the Madden-Julian Oscillation can be used to improve the sub-seasonal to seasonal forecasts of precipitation which can be used to improve water management.

ACKNOWLEDGEMENT

I would like to express my gratitude to everyone who has helped make this possible. To my advisers Balaji and Tom, I wouldn't be here without you, and your serendipitous meeting. Your advice, both professional and personal, and your assistance made this dissertation possible. To the rest of my committee, Edie, Ben, and Jean-Francois, and to everyone else who has taken time to answer my emails and phone calls, your guidance and help has been invaluable. To my family, Mom, Dad, Amy, you have always pushed me to follow my own path, and always been there when I needed you. Thank you and I love you. And finally to my friends, my Boulder family, you have always been there to offer encouragement, and much needed distractions. You're awesome and I love you all. Thank you!

CONTENTS

TABLES

FIGURES

1 INTRODUCTION AND BACKGROUND

1.1 INTRODUCTION

There has been an increasing interest in and need for the application of climate information to address societal issues. Partially arising from the increasing confidence in anthropogenic climate change (Stocker et al. 2013) and partially from increasingly complex and resource-limited conditions, climate impacts are now considered in fields as diverse as natural hazards, public health, and agriculture. The research community has begun to address the practical concerns associated with this interdisciplinary work, standardizing methodologies, developing new tools, and investigating the repurposing of existing tools (Thornton et al. 2009; Nardone et al. 2010; Rosenzweig et al. 2013). This research looks to show the utility of existing data analysis and statistical tools in providing actionable information to decision makers. Though this information is relevant in a wide variety of locations and conditions, the focus of this research has been on producing information for applications in the developing world. The international community has made the impacts of climate change on the developing world a focus in the recently released Sustainable Development Goals (SDG) (United Nations General Assembly 2015). The SDG highlight a need for local, targeted information for evaluating the impacts of climate change, identifying potential solutions, and providing information that allows for mitigation of climate change impacts. Of the 17 high-level goals, three are directly related to this research. These include:

"2.4: By 2030, ensure sustainable food production systems and implement resilient agricultural practices that increase productivity and production, that help maintain ecosystems, that strengthen capacity for adaptation to climate change, extreme weather, drought, flooding and other disasters"

"13.2: Integrate climate change measures into national policies, strategies and planning"

"13b: Promote mechanisms for raising capacity for effective climate change-related planning and management"

Climate variability at multiple timescales can have an outsized impact in the developing world given the often limited infrastructure, and weaker social and economic institutions (Field et al. 2014). There is a need for tools and analyses that translate climate information at coarse space and time scales to local scales where decisions of resource management are made. Development of these tools requires an understanding of the spatial and temporal variability of the underlying climate variables affecting the processes of interest. This research presents two case studies that follow the following methodology:

a) examine and quantify the seasonal and interseasonal variability of the climate variables of interest
b) identify links between these climate variables and large-scale drivers including local phenomena and teleconnections
c) select a measure that quantifies the impact of interest, and examine how it is affected by climate variability
d) use the understanding of climate variability and the impact measure to develop current conditions, forecasts, or other products targeted at planners and decision-makers

The two case studies are located within the monsoon regions of Africa, one in West Africa and one in East Africa (Figure 1-1). The application of the tools and information produced are targeted at local issues – livestock heat stress in West Africa where agriculture is a major livelihood, and water resources in East Africa where there has been large-scale and ongoing development of hydropower projects.

The West and East African monsoons, WAM and EAM, respectively, dominate the local climate during the monsoon season of June through September, both have demonstrated connections to large-scale climate features, and both have exhibited large interannual variability.

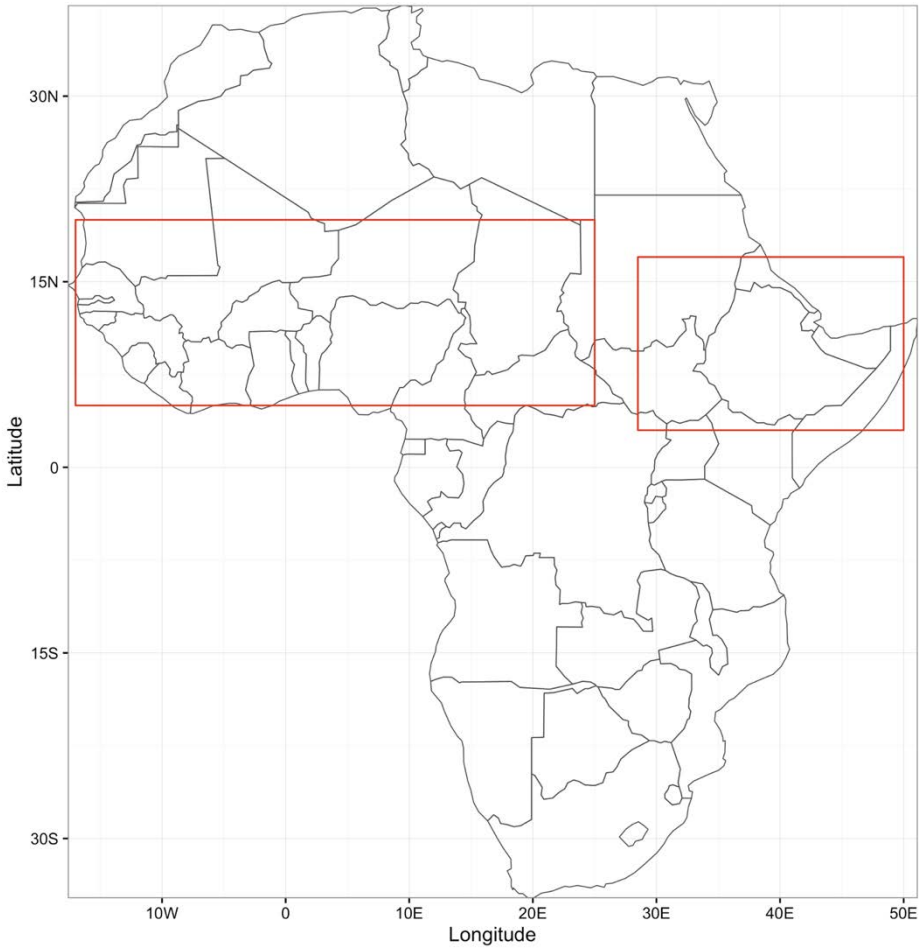

Figure 1-1: West African Monsoon (left) and East African Monsoon (right) Study Regions (shown in red boxes)

1.1.1 WEST AFRICA

The focus for the first case study is the Sahelo-Sudanian region of West Africa, where agricultural production in aggregate supports 60% of the active population and is responsible for 40% of the gross national product (Battisti and

Naylor 2009). Understanding the variability and forecasting of measures related to livestock health and production can improve livestock management, and have a large impact on one of the primary livelihoods in this region. Driven by the seasonal migration of rainfall and productive grassland, as well as the seasonal cycle of livestock disease, a pastoral transhumance system has developed. Livestock herders migrate north and south throughout the year: north during the monsoon period to take advantage of seasonal grasslands, and escape the higher risk of livestock disease in the south, and south following the start of the dry season in the north. The seasonal climate of the region is controlled by the West African Monsoon (WAM), a low-level southwesterly flow existing during the boreal summer. The WAM is largely driven by a pressure gradient existing between two centers of action, the South Atlantic High, and the West African Heat Low (WAHL). The relative strength of these centers determines the strength of monsoon winds, and indirectly controls moisture advection onshore. Seasonal strength of the WAM impacts both the magnitude and location of summer rainfall, and also influences seasonal humidity and temperature. The WAM transition periods, termed 'monsoon onset' and 'monsoon retreat' can be identified by the seasonal migration of the Intertropical Convergence Zone (ITCZ) and the Intertropical Front (ITF), the later the interface between monsoon and Harmattan winds. The timing of onset and retreat also have a large influence on seasonal rainfall, humidity, and temperature.

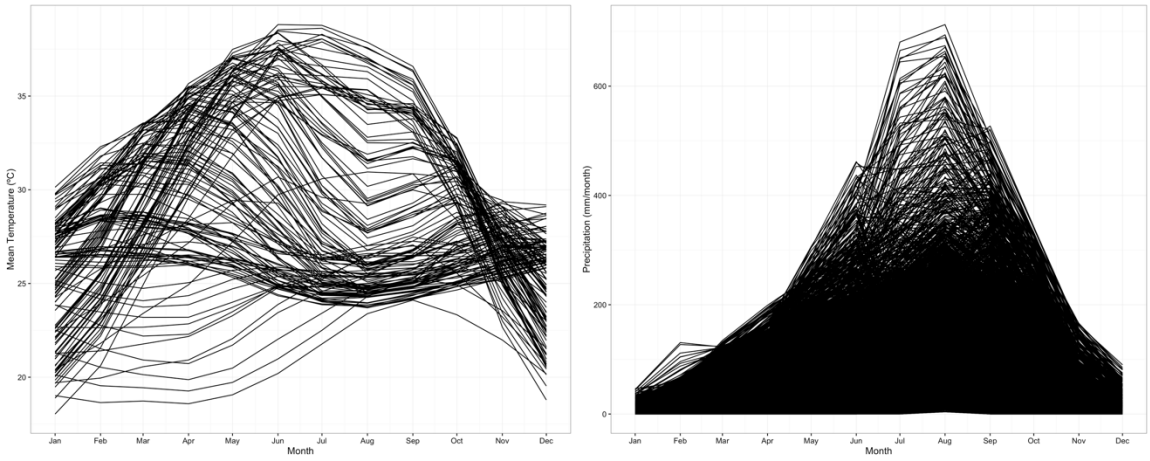

Figure 1-2: West Africa monthly surface temperature climatology (left) and precipitation climatology (right) for each observation location in the region (from the HadISD dataset)

Figure 1-2 shows the monthly climatology of temperature and rainfall at observation locations across the region. The inland Sahelo-Sudanian region of West Africa receives majority of the annual total precipitation during the June through September summer season. This precipitation has experienced large interannual variability over the 20th century – a wet period in the 1950s and 1960s proceeded by the 1970s-2000 Sahel drought (Dai et al. 2004; Nicholson 2005; Lebel and Ali 2009; Dieppois et al. 2014). Variability has been linked to a variety of causes, but is dominated by SST teleconnections including the Gulf of Guinea, the tropical Atlantic Ocean, the Indian Ocean, and the tropical Pacific Ocean (ENSO) (Losada et al. 2012). These SST teleconnections influence precipitation through two primary mechanisms: moisture transport into the region, and dynamical connections that promote convection *Nicholson*, 2013]. Quantifying the variability of these teleconnections as well as the subseasonal variability can be leveraged to improve forecasting. This can also illuminate how modern-day SST warming has impacted these teleconnections and ultimatly precipitation dynamics. Changes in local precipitation directly impact the local population in a number of ways including heat stress expericenced by livestock. As a case study for this region, we investigate

livestock heat stress and tools that can use this improved understanding of subseasonal and interannual climate variability to provide better information for decision-makers.

1.1.2 LIVESTOCK HEAT STRESS

The focus of our first case study (West Africa) is on livestock heat stress. Livestock heat stress is the physiological response to direct environmental exposure and manifests as an increase in an animal's core temperature, leading to decreases in metabolism, decreases in production, and in severe cases, death. It can also impact an animal's ability to successfully reproduce. Quantifying heat stress burden and risk in the developing world has been limited, with little understanding of the susceptibility of native breeds to heat stress. This lack of knowledge has been highlighted by the International Livestock Research Institute (ILRI), who state: "the impacts of increased frequencies of extreme heat stress on existing livestock breeds are not known, nor do we know if there are critical thresholds in the relationship between heat stress and physiological impacts" Easterling and Apps (2005) state: "Confidence in the ability of livestock producers to adapt their herds to the physiological stresses of climate change is difficult to judge. The absence of physiologically based animal models with well-developed climate components suggests a major methodological void." Additional complexity is introduced when considering a broad region such as sub-Saharan Africa where there is a large heterogeneity in livestock systems in terms of breeds, methodology, and adaptive capacity. Adaptation and selective breeding have produced cattle breeds with greater resistance to local conditions including climate and disease. The resilience of these breeds to heat stress has not been well quantified; with the threshold values commonly used with heat stress indices developed for intensively managed cattle in the developed world. The lack of complete, long-term climate data has affected the ability to establish baseline heat stress burden and evaluate changes.

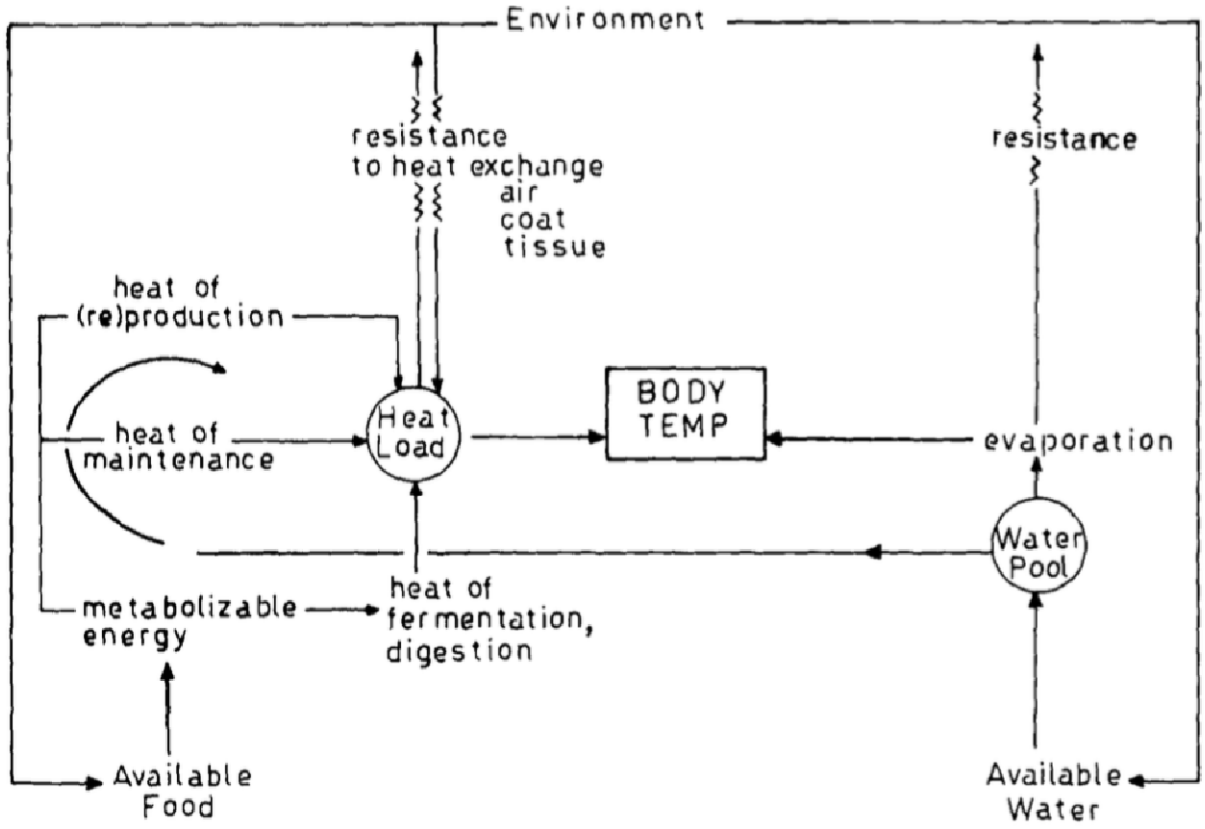

Figure 1-3: Conceptual livestock heat stress model showing the external and internal components (figure from Finch 1986).

Heat stress is a function of temperature, humidity, solar radiation, and windspeed directly, and rainfall indirectly. Both index-based and physiological-model based approaches have been taken to quantify heat stress and are described in detail in Chapter 3. These approaches all require daily weather sequences, which are conditioned by the monsoon variability. This work demonstrates the utility of a stochastic weather generator (SWG) incorporating large-scale climate information to produce statistically consistent daily weather sequences. Consequently, heat stress indices are generated, for livestock health management. The SWG offers a flexible tool to generate daily weather sequences at seasonal and multi-year time scales. The use of a SWG for simulating heat stress attributes is unique and one of the first applications to this field.

1.1.3 EAST AFRICA

The focus for the second case study is on the northern East Africa. The Kiremt season, June through September, is the primary precipitation season for East Africa north of about 7ºN. This region is the headwaters of two major rivers, the Blue Nile, and the Omo Gibe, as well as a major agricultural region. Decreases in precipitation occurred in concert with decreases in the West African Sahel during the 1970s and 1980s, however, while there has been a 'recovery' in the Sahel, precipitation has continued to decrease in the recent 1990-present period (Williams et al. 2012). Variability in the influence of large-scale teleconnections on Sahel precipitation over the past century have been noted in Broman et al., submitted. Similar to West Africa, precipitation variability in this region can be broadly linked to two mechanisms: moisture transport into the region, and the behavior of local dynamical features promoting convection. Local circulation features are controlled by the East African Monsoon (EAM) which establishes itself in the boreal spring and extends through the summer. At a seasonal timescale, remote teleconnections with ENSO, the tropical Indian Ocean, and the tropical Atlantic Ocean influence moisture transport, the behavior of the EAM, and thus precipitation variability. At subseasonal scales, precipitation variability has been linked to the Madden-Julian Oscillation (MJO). This variability, both interannually, and interseasonally, has impacts on local water resource management and agriculture. An understanding of this variability, connections to large-scale teleconnections, connections to the MJO, and how these connections have varied over time provide a basis for improving forecasts. Training statistical forecast models understanding the variability in the covariates, or post-processing numerical weather prediction models (NWP) offer two ways of incorporating in this information. The next section describes an approach for post-processing NWP forecasts at the subseasonal to seasonal (S2S) timescale, 0-60 days to improve forecast skill.

1.1.4 PRECIPITATION FORECASTING

The focus of our second case study is on improving S2S precipitation forecasts targeting water management and agriculture. Improvements in forecasts have the potential to improve management decisions in the developing world (Webster 2013). The S2S timescale has seen increased interest in recent years given the timescales of decisions (Robertson et al. 2015).

There have been several attempts to develop skillful seasonal precipitation forecasts for the Kiremt season. Statistical regression approaches have demonstrated spatially-dependent skill at short lead times using large-scale ocean and atmospheric features. (Gissila et al. 2004; Korecha and Barnston 2007; Block and Rajagopalan 2007; Diro et al. 2011; Nicholson 2014). Predictability was strongly linked to the summer ENSO state, as well as discerning intensifying versus decaying behavior in the spring. Skillful forecasts were limited to a two month lead time by the 'predictability barrier' of ENSO (Torrence and Webster 1998), rapid changes in circulation during the March-May season, and the sensitive of the tropical Pacific and Indian Ocean to perturbations. Seasonal skill in numerical weather prediction (NWP) forecasts is limited deterministic skill beyond a two-month lead time (Kirtman et al. 2014). Combining NWP forecasts at the S2S timescale, with post-processing approaches that can incorporate in information about teleconnections (e.g. ENSO) and subseasonal controls (e.g. MJO) offer the ability to improve forecast skill.

1.1.5 DISSERTATION STRUCTURE

The structure of this proposal is as follows. Chapter 2 presents an overview of the large-scale climate features controlling precipitation variability in the West African Sahel. Variability in large-scale teleconnections over the past century is investigated using composite analysis and Bayesian dynamical linear models (BDLM). These analyses offer insights into internannual and interseasonal precipitation variability. Chapter 3 demonstrates the utility of a stochastic weather generator to produce statistically consistent weather sequences for use in

evaluating climate-related impacts. Given the social and economic importance of livestock, and the limited investigation, we use the weather generator to examine livestock heat stress impacts in West Africa. We show the ability to improve weather generator output to match observed statistics when conditioned with large-scale climate. Spatial modeling of both weather generator-derived heat stress estimates as well as estimates obtained directly from observations allow point measurements to be extended over the entire study region. When coupled with the monsoon diagnostic information from Chapter 2, seasonal predictive tools can be developed to provide planners and decision-makers information to mitigate the impacts of potential heat stress events. Chapter 4 repeats the same analysis as Chapter 2 except for East Africa. In addition to large-scale teleconnections, we also examine the influence of the MJO on precipitation variability at the subseasonal timescale. Chapter 5 evaluates the skill of the CFSv2 precipitation forecasts for the summer monsoon season in East Africa. We examine the impact of quantile regression (QR) post-processing directly and QR incorporating in covariates. Finally, Chapter 6 provides a review of the findings from Chapters 2-5.

2 SPATIAL AND TEMPORAL VARIABILITY OF SAHELIAN WEST AFRICA PRECIPITATION AND LARGE-SCALE TELECONNECTIONS

2.1 INTRODUCTION

The late 20th century Sahel drought, lasting from the 1970s through the 1990s (Dai et al. 2004; Nicholson 2005; Lebel and Ali 2009; Dieppois et al. 2014) and the wet period preceding present a clear illustration of the variability of precipitation in the inland Sahelo-Sudanian region of West Africa. These two 'epochs' along with a modern period epoch beginning at the turn of the 21st century represent distinct shifts in the underlying drivers of precipitation in the region. This epochal variability has been attributed to a variety of causes including changes to the land surface producing an albedo feedback loop (Charney 1975; Nicholson 2000), and changes to sea-surface temperature (SST) driven teleconnections (Folland et al. 1986; Palmer 1986; Fontaine and Janicot 1998; Giannini et al. 2003; Lu and T.L. 2005; Koster et al. 2004; Xue et al. 2016). Via their effect on the main dynamical features controlling precipitation variability – the occurrence of Mesoscale Convective Systems (MCSs), and moisture availability (Nicholson 2013) – variability in these teleconnections are now thought to be the dominant driver of summer precipitation variability (Losada et al. 2012). MCSs are responsible for large-scale precipitation in the region, upwards of 90% of the seasonal total (Lebel et al. 2003; Nesbitt et al. 2003; Mohr and Thorncroft 2006). SST teleconnections influence MCSs indirectly through the strength and location of the mid-level African Easterly Jet (AEJ) and high-level Tropical Easterly Jet (TEJ). MCSs are organized in part by African Easterly Waves whose strength and propagation are in turn controlled by the AEJ and TEJ (Jackson et al. 2009). Through this chain, northward shifts in the center of the AEJ have been observed in wet years, extending as far north as 20º in some years. Conversely, a more southerly AEJ has been observed in some dry years. A similar correspondence is seen with the strength of the TEJ: a stronger TEJ is associated with increases in precipitation, and a weaker TEJ associated with decreases in precipitation (Hulme and Tosdevin 1989;

Nicholson and Grist 2001; Nicholson 2013). SST teleconnections have been linked to variability in these two jets; the effect of different SST regions are discussed in detail below. Moisture availability is drive by the West African Monsoon (WAM), a low-level flow that advects moisture inland from the Gulf of Guinea and Atlantic Ocean. The strength of the cross-equator pressure gradient driving this flow affects the availability of moisture in the Sahelo-Sudanian region, and is also controlled by SSTs, both directly, and through teleconnections.

Four SST regions have been linked through the above two dynamical features to precipitation variability – the tropical Pacific Ocean, the tropical Indian Ocean, the tropical Atlantic Ocean, and the Gulf of Guinea. The tropical Pacific influences precipitation through the El Nino Southern Oscillation (ENSO). A more northerly AEJ is seen during La Nina, and a more southerly AEJ during El Nino. A weakening of the TEJ is also observed during El Nino conditions (Chen and van Loon 1987; Nicholson 2013). There is some debate about the influence of the tropical Indian Ocean. A modeling study (Hagos and Cook 2008) indicated that a warming Indian Ocean could have produced a region of subsidence over the Sahelo-Sudanian region, blocking monsoon moisture transport, and contributing to the decreases in precipitation seen in the 1980s and 1990s. The same modeling study suggest that continued warming would have shifted this subsidence westward over the Atlantic, and could explain the 'recovery' of precipitation in the early 21st century. A similar modeling study shows a more local influence of the Indian Ocean focused on the western Sahel (Bader and Latif 2003). Conversely, statistical analyses of observed precipitation indicate a relative minor influence of the Indian Ocean (Janicot et al. 2009). This study describes an atmospheric teleconnection between the Indian Monsoon and West Africa that increase low-level westerlies and moisture transport, but whose influence is muted compared to the influence of other SST regions. In the Atlantic, the relative difference in temperatures between the northern and southern tropical regions (tropical Atlantic dipole) affects the pressure gradient that drives monsoon moisture transport. The Gulf of Guinea plays a similar role, where increases in SSTs weaken this pressure gradient and therefore moisture transport

(Vizy and Cook 2002; Fontaine and Louvet 2006; Caniaux et al. 2011). Historically there has been a prominent dipole pattern between precipitation in the Guinea Coast region and precipitation in the Sahelo-Sudanian region (Nicholson et al. 1998; Nicholson 2013; Losada et al. 2012). This pattern appears to have weakened or disappeared since the start of the Sahel drought in the 1970s (Nicholson 2013; Losada et al. 2012) suggesting changes in the relationship between precipitation variability, the Gulf of Guinea, and the tropical Atlantic dipole.

Global SSTs appear to be the dominant driver of interannual and interdecadal variability of precipitation in this region. However, there does not appear to be a single clear mechanism. Rather, it is the combined behavior of all four regions that control variability. Similarly, from work on the Indian Monsoon teleconnections appear to influence precipitation unevenly at subseasonal timescales [Gill et al. 2015]. Examining this behavior in other monsoon regions could suggest broader physical link between the Indian and West African monsoons, or dynamical controls responsible for this uneven behavior broadly across monsoon regions.

To date, there has not been any comprehensive study of the variability of these teleconnections at either seasonal or subseasonal time scales. These gaps motivate this research. In this paper, we seek to quantify the relative strengths of SST teleconnections over the 20th and early 21st centuries as well as investigate their subseasonal variability. Using composite analysis and Bayesian Dynamical Linear modeling (BLDM), we investigate the space-time variability of West African summer monsoon precipitation and associated SST teleconnections.

To connect this analysis with the existing understanding of precipitation variability we consider the full period (1901-2014), and three precipitation epochs separately. The paper is organized as follows: a description of the datasets, regional climatology and methods used are first presented, results from the analysis of the full period that describe the long-term variations in precipitation and SST teleconnections are next presented, followed by the results from epochal analysis

examining the dynamics of this variability during the different periods; concluding with summary and discussion of the results.

2.2 DATA AND METHODS

Gridded precipitation data were obtained from the CRU TS3.23 dataset which provides monthly coverage for a 114-year period from 1901 to 2014 at a 0.5°x0.5° resolution (Harris et al. 2014). Summer season (June-September) average and subseasonal average – June (early season); July-August (peak season) and September (late season) – precipitation values were computed from this dataset. Global climate fields – winds, geopotential heights, sea level pressure, and precipitable water – were obtained from the NCEP/NCAR Reanalysis 1 (NNR) (Kalnay et al. 1996). Four SST index timeseries are used – Nino 3,4, Indian Ocean Dipole Mode Index (IOD-DMI), the Tropical Atlantic Index (TAI), and Gulf of Guinea Index (GGI) – and provide timeseries of standardized SST anomalies over their respective regions. Nino 3,4 index, bounded by 120°W-170°W and 5°S- 5°N (Trenberth 1997) was obtained from NOAA Earth Systems Research Laboratory (ESRL) (esrl.noaa.gov). The remaining indices were computed from the Hadley-OI merged SST dataset (Hurrell et al. 2008) as follows: the IOD-DMI as the temperature gradient between the East (90°E-110°E and 10°S-Equator) and West (50°E-70°E and 10°S-10°N) Indian Ocean (Saji et al. 1999), the Tropical Atlantic Index (TAI) as the temperature gradient between the TSA, bounded by 5.5°N-23.5°N and 57.5°W-15°W, and TNA bounded by 5.5°N-23.5°N and 57.5°W-15°W (Enfield et al. 1999), and the GGI bounded by 12ºW and 10ºE and 5ºS to 8ºN.

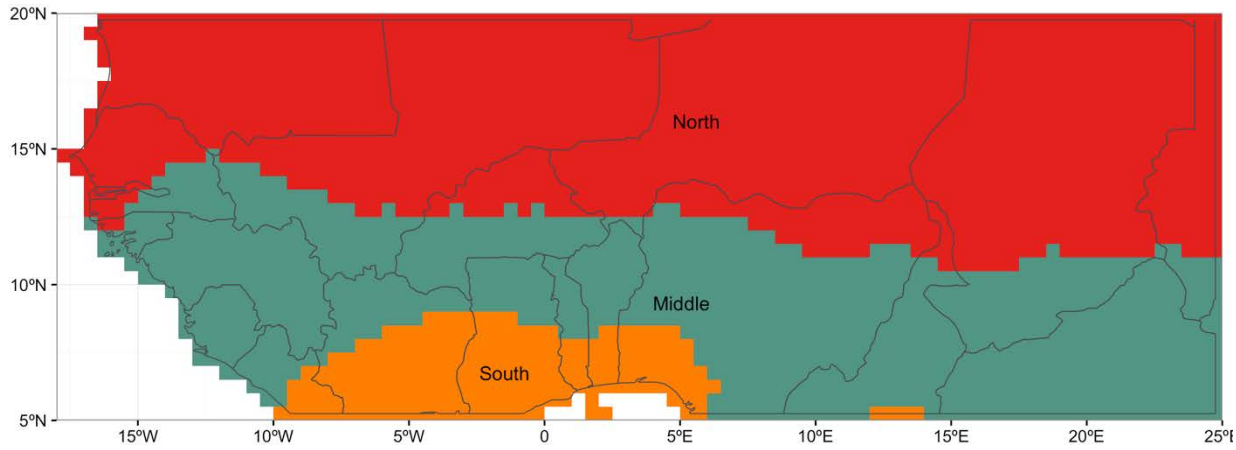

Figure 2-1 West Africa precipitation clusters, using a k-Means clustering approach on annual climatology and the CRU TS3.23 precipitation dataset

To identify coherent precipitation regions, a K-means cluster analysis (Scott and Knott 1974) was performed on the annual climatology of precipitation. Three distinct clusters emerged and are shown in Figure 2-1. These are referred to as South', 'Middle', and 'North', with the Middle and North cluster areas largely corresponding to the Sahel region. Therefore, we focus on these two cluster areas. For each of these two cluster areas, anomalous precipitation years were identified using a cluster-average standardized anomaly timeseries (1950-1999 climatology) with low years <= -1 and high years >= 1. Anomalous years were identified based on the seasonal and the three subseasonal average precipitation timeseries separately. Patterns of large-scale climate fields is investigated through composite analysis – average of each field in low and high years. We performed composite analysis of SST anomalies, low-level geopotential height (GPH) (850hPa), low (925hPa), mid (600hPa), and high (200hPa) level winds, and total-column precipitable water (precipitable water).

To examine the changes in strength of large-scale climate features, a Bayesian Dynamical Linear Model (BDLM) was used. In traditional linear regression, models the coefficients are static, while in BDLM they vary as a function time thus enabling to capture changing relationships, i.e., nonstationarity, BDLM (West and Harrison 1997; Petris 2009). This has been applied in modeling temporal

variability of teleconnection of Indian summer monsoon and SSTs from Pacfic and Indian Ocean (Maity and Nagesh Kumar 2006; Krishnaswamy et al. 2015), and similarly with rainfall over Indonesia [*Yanto et al.* 2016]. Its use led to interesting insights and better performance than the traditional regression methods, showing for the Indian case a strengthening influence of Indian Ocean SSTs on the summer monsoon and extreme precipitation events in recent decades and the weakening influence of ENSO. In the case of Indonesia, it documented the temporal variability of the relationship between ENSO and precipitation. Recently, this has also been applied to model the variability of sea level in Northern Atlantic (Kennigson et al., 2018).

In this method, the regression coefficients vary with time, unlike traditional regression where the coefficients remain fixed. In BDLM the time series is considered as the output of a dynamical system perturbed by random disturbances (i.e., noise) – thus, considered a nonstationary evolution. The general form of the model is represented as:

$$y_t=F_t\ x_t+v_t\ v_t\sim N(0,V_t)$$
$$x_t=G_t\ x\_(t-1)+w_t\ w_t\sim N(0,W_t)$$

where y_t is a dependent variable (e.g. the leading PC), x_t is a vector of independent variables (e.g., Nino 3,4, TAI, etc.), t denotes time, F_t is an observation operator that transforms the model states into observations (e.g. regression coefficient) and G_t is the linear system operator (that models states of the independent variables). Both observation and system equations can have additive Gaussian errors with covariance matrices V_t and W_t.

The posterior predictive distribution of model coefficients θ_t at each time t is computed from the prior distribution of coefficients at time step t-1. Using Bayes theorem, the probability of the data yt conditional on the model parameters at time t is defined as:

$$P(\theta_t \mid y_t) \propto P(y\ y_t \mid \theta_t) P(\theta_t \mid y_{t-1})$$

The model coefficients θ consist of PCs, Nino 3,4, PDO, the regression coefficients and the variances of the Gaussian errors in the above equations. In the initial step, traditional linear regression is applied to compute regression coefficient with normally distributed mean and variance – i.e. $\theta_0 \approx N(m_0, C_0)$ where m_0 and C_0 are the mean vector and the variance – covariance matrix of the regression parameters. Using a Kalman filtering approach together with Markov Chain Monte Carlo simulation approach, posterior distributions are estimated at any time t. The posterior distribution is then employed to generate the Bayesian confidence intervals (Petris 2009).

The BDLM returns transient regression coefficients for each covariate included, in this case the four SST indices, Nino 3,4, the TAI, the DMI, and the GGI. These time-varying coefficients can be compared to a fixed regression coefficient and its 5-95% confidence bounds (CI), to identify epochal shifts. We used these shifts to define three epochs and repeat the above composite analysis.

2.3 RESULTS

As mentioned we focus on the Sahelian region – i.e., northern cluster. Results from the composite analysis to understand the seasonal and subseasonal teleconnections with large-scale climate variables are first presented followed by the BDLM analysis for the temporal variability.

2.3.1 SEASONAL AND SUBSEASONAL COMPOSITE ANALYSIS

Composites of precipitation of high (i.e. wet) and low (i.e. dry) years for the full season and the three subseasonal periods are shown in Figure 2-2. As expected, the northern cluster – Sahelian West Africa – is uniformly wet and dry during the respective years (top row) and they are consistent across the subseasons. This wet / dry anomaly pattern can be seen outside the Sahel, in northern East Africa, and

into the Indian subcontinent, especially during late season. During peak season the positive and negative anomalies clearly extend through to East Africa, and weakly during the early and late seasons. This indicates subseasonal heterogeneity in the spatial variability of precipitation and potentially teleconnections that would not be apparent from a seasonal analysis.

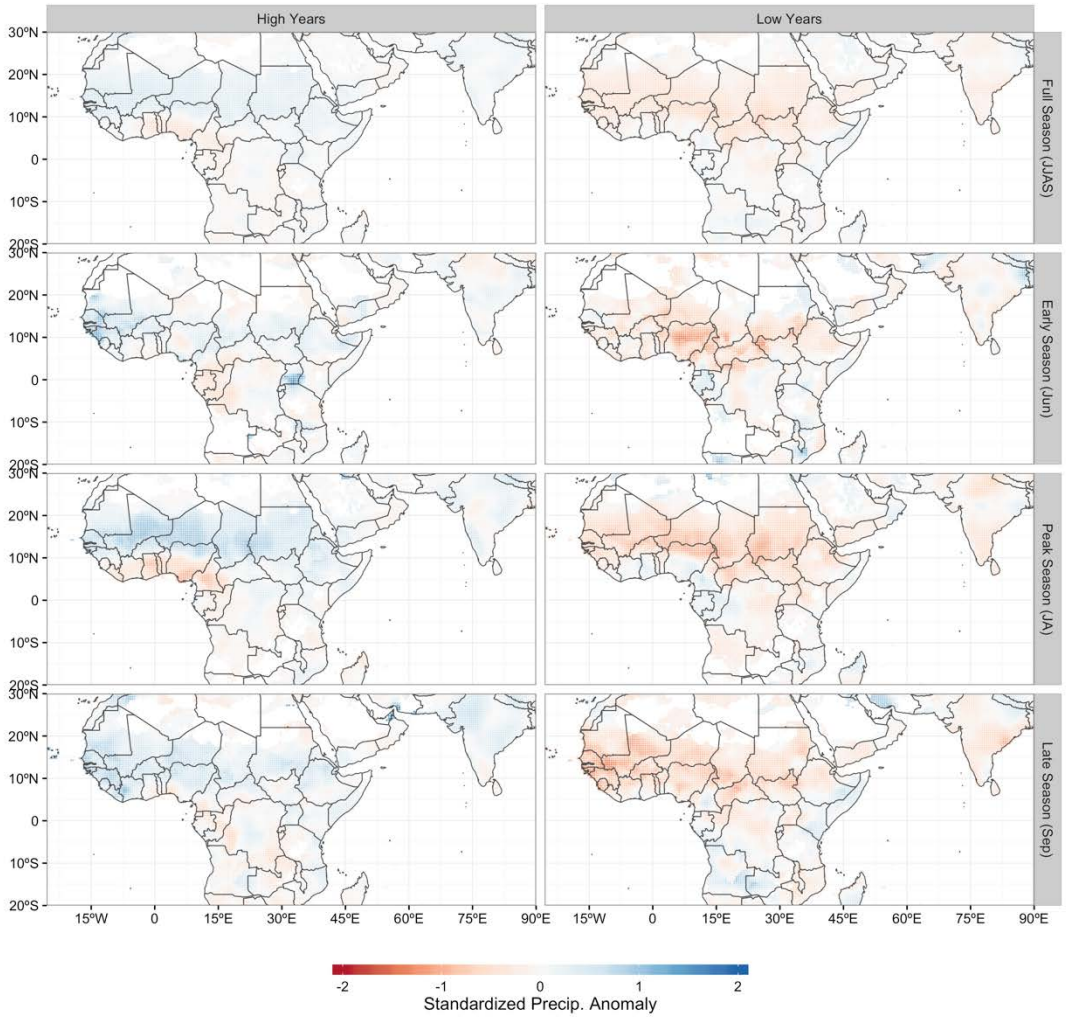

Figure 2-2 Composites of average precipitation corresponding to high (left column) and low (right column) precipitation years in the northern cluster based on the full record (1901-2012). Top row shows the composites for full season, June-September, 2nd row the early season, June, 3rd row peak season, July-August, and bottom row, late season, September.

In order to identify the large-scale climate features accompanying these precipitation regimes composites of SSTs, 925hPa, 600hPa, and 200hPa level winds, precipitable water, and 850hPa GPH were created for the full season and sub-seasons.

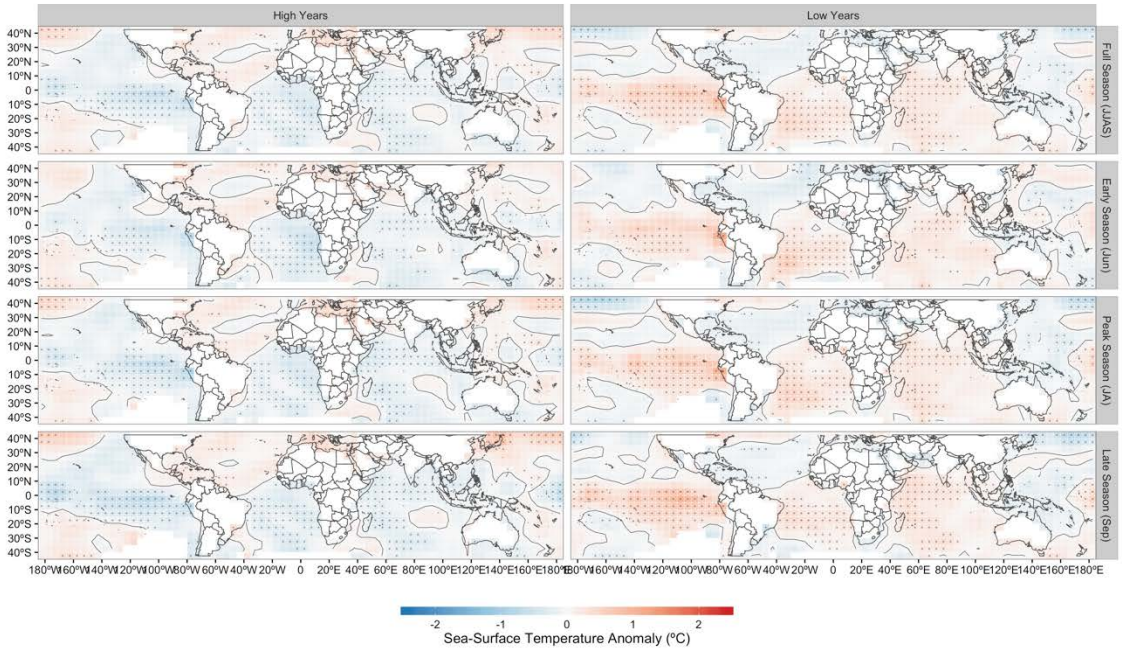

Figure 2-3 Same as Figure 2-2 but composites of SST anomalies. Statistically significant regions are stippled.

The SST anomaly composites for the high and low years of the northern cluster are shown in Figure 2-3. The statistically significant regions are in the tropics of Atlantic, Pacific and Indian Ocean basins and they are consistent across the sub-seasons. The Atlantic dipole behavior, with a cool southern tropical Atlantic and warm northern tropical Atlantic during high precipitation years and the converse during low precipitation years, over the entire season is evident (Figure 2-3**,** top row). The southern tropical Atlantic region shows stronger SST anomalies compared to the northern tropical Atlantic and is the case during the sub-periods. Furthermore, the cooler SST anomalies in southern tropical Atlantic, especially closer to the African coast, during wet years are stronger compared to the warmer SST anomalies during the low years (top row). This is also consistent during all the

subseasonal periods. This asymmetry is noteworthy. Wetter years are accompanied by cooler central and eastern tropical Pacific, reminiscent of La Nina conditions, vice-versa for drier years, reminiscent of El Nino conditions (Hulme and Tosdevin 1989; Nicholson and Grist 2001; Nicholson 2013). The signal in the Indian Ocean is weaker in comparison.

Composites of anomalies of 925hPa winds, precipitable water, and 850hPa GPH are shown in Figure 2-4. The 925hPa winds correspond to the monsoonal jet in the region (Nicholson 2013) and the 850hPa heights are selected to avoid surface friction. During high years the low-level southwesterly monsoon winds from the Atlantic are strong. This is also accompanied by strong northeasterlies coming from the Arabian Peninsula and northern Africa, consistent with anomalous low pressure over the region (Figure 2-4 contours). There are positive precipitable water anomalies (Figure 2-4, green colors) over central and western Africa (centered over Chad and Niger) during high precipitation years.

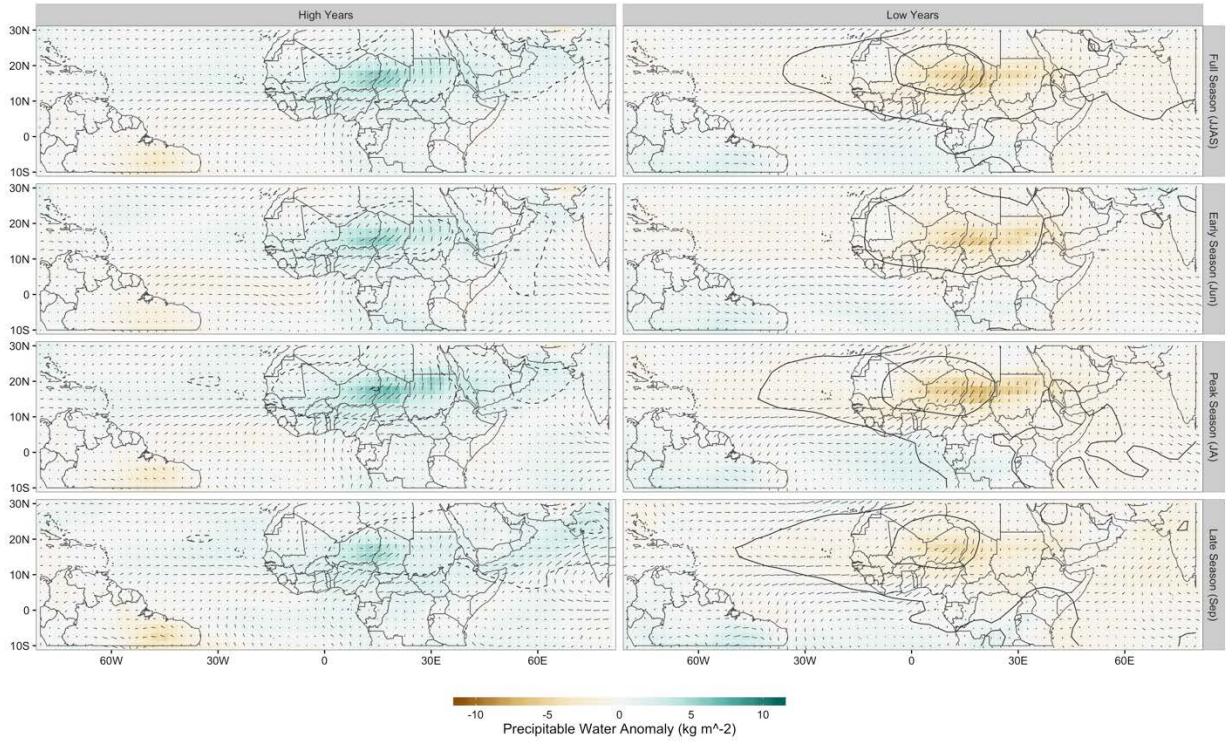

Figure 2-4 Same as Figure 2-2, but for average 925hPa winds (arrows), pressure (contours) and precipitable water (colors).

This feature is seen for the full season and all three subseasonal periods, being most pronounced during the June early season and the July-August peak season. During the September late season, the convergence of winds and moisture is relatively weaker (Figure 2-4, bottom panel). However, notice that the Indian monsoon jet, extending over Arabian Sea and on to Indian subcontinent, is strong in low years (bottom-left panel), which corresponds with the late season of the Indian summer monsoon. This was noticed in Gill et al., 2015b during La Nina years. Thus, during wet years, tropical Atlantic and Pacific are anomalously cooler, with a strong low pressure center formed over the Sahelian region, enabling convergence of low-level winds and moisture from the southwest and northeast. Furthermore, the atmospheric teleconnections are stronger during early and peak season. In low precipitation years the low-level south easterlies are virtually absent, and the tropical Oceans are warmer with El Nino conditions in the Pacific (Chen and van

Loon 1987; Nicholson 2013). These anomalies are also consistent across the sub-seasons.

2.3.2 TEMPORAL VARIABILITY OF TELECONNECTIONS

The BLDM was fitted to the summer season precipitation in the north cluster with the contemporaneous large-scale SST teleconnection indices and time varying coefficients for the intercept, Nino 3,4 TAI, DMI, and GGI are shown in Figure 2-5. Coefficients from a stationary or static (i.e. standard linear regression) regression and their 90% confidence intervals are shown as horizon lines in these figures. Periods where the BLDM time-varying regression coefficient deviate outside the confidence bounds indicate it is significantly different and can be interpreted as changes in the teleconnection strength.

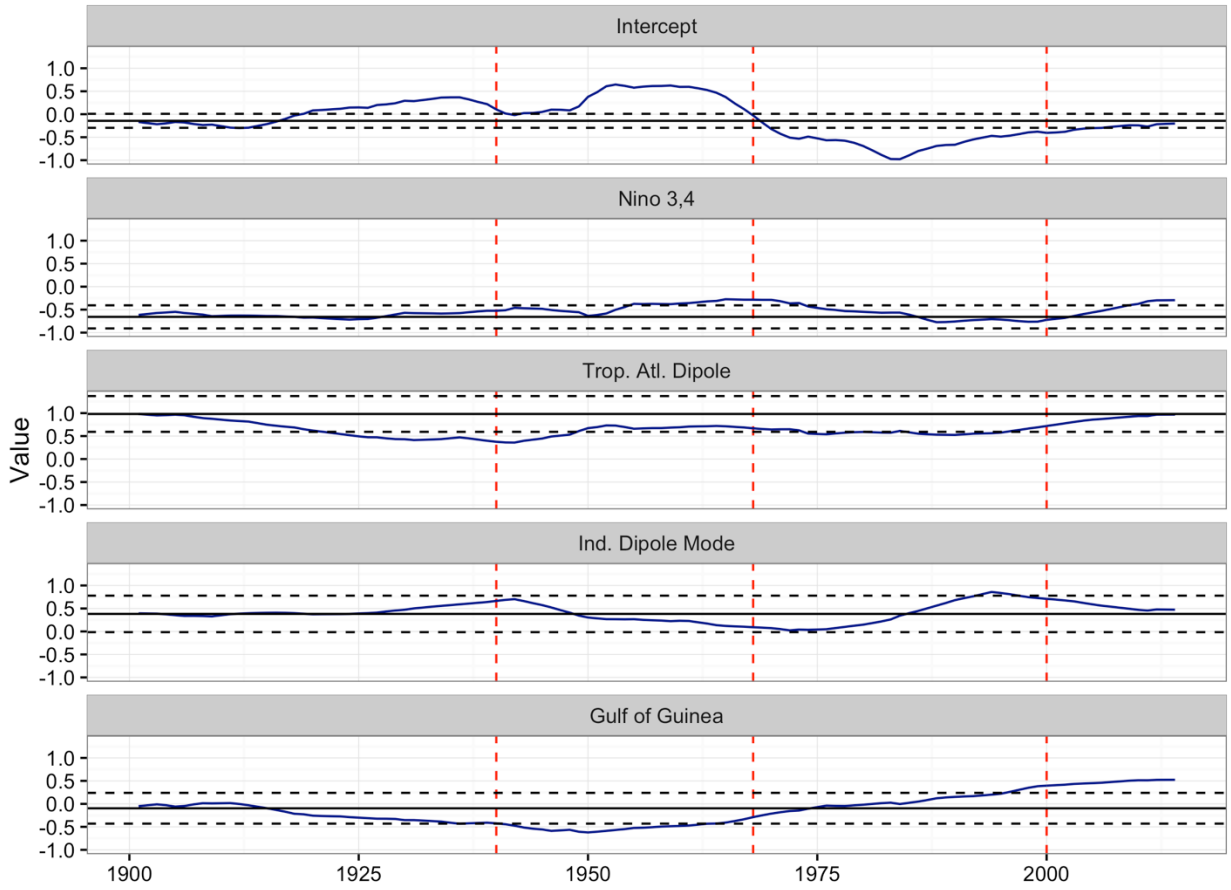

Figure 2-5 Time varying regression coefficients – intercept (top panel), Nino 3,4 (2nd panel), TAI (3rd panel), DMI (4th panel), and GGI (bottom panel) – from BDLM for north cluster precipitation (blue lines). Coefficients from stationary linear regression are shown in black with 95% confidence intervals (dashed lines). Red lines show the boundaries of the defined epochs – wet period (1945-1968), drought (1969-2000), and modern period (2001-2014).

The intercept, which can be viewed as the mean value of precipitation anomaly, shows an increasing trend through 1960s and then a significant decline through mid-1990s, before increasing. This is consistent with the Sahelian precipitation variability reported in literature (Dai et al. 2004; Nicholson 2005; Lebel and Ali 2009; Dieppois et al. 2014). The coefficient of Nino 3,4 has remained negative over the entire 114 years, it weakened (i.e. value moving towards zero) starting in the 1930s before returning to the static regression value in the 1970s. Around 2000 this

relationship has again begun to weaken as seen during the mid-century. This epochal variability with Nino 3,4 is consistent with the epochal variations of Indian Monsoon-ENSO teleconnections which has been shown to weaken in recent decades (Kumar et al. 1999). The TNA has exhibited a positive relationship with precipitation, and the transient regressions have tended to remain within the confidence intervals of the static regression values. However, it has weakened during 1900 through 1940s and increasing since mid-1990s. Interestingly, this appears to be opposite to that of what was seen with Nino 3,4, in that, when the strength of ENSO teleconnections is weakening, the Atlantic teleconnections appear to be strengthening (Losada et al. 2012)

The static regression indicates a positive relationship with IOD-DMI, while the BDLM which shows weakening around 1950, before strengthening around the mid-1970s, and declining in recent times. However, it remains near the static regression value throughout. This epochal variability in Indian Ocean teleconnections has also been seen with Indian monsoon (Krishnaswamy et al. 2015) in that in recent decades the strength of Indian Ocean teleconnection with the Indian monsoon is stronger when the ENSO teleconnection is weakening, which is consistent here. The Gulf of Guinea SSTs show no relationship through mid-century when it exhibited negative relationship and since the late 1990s there is a shift to a positive relationship.

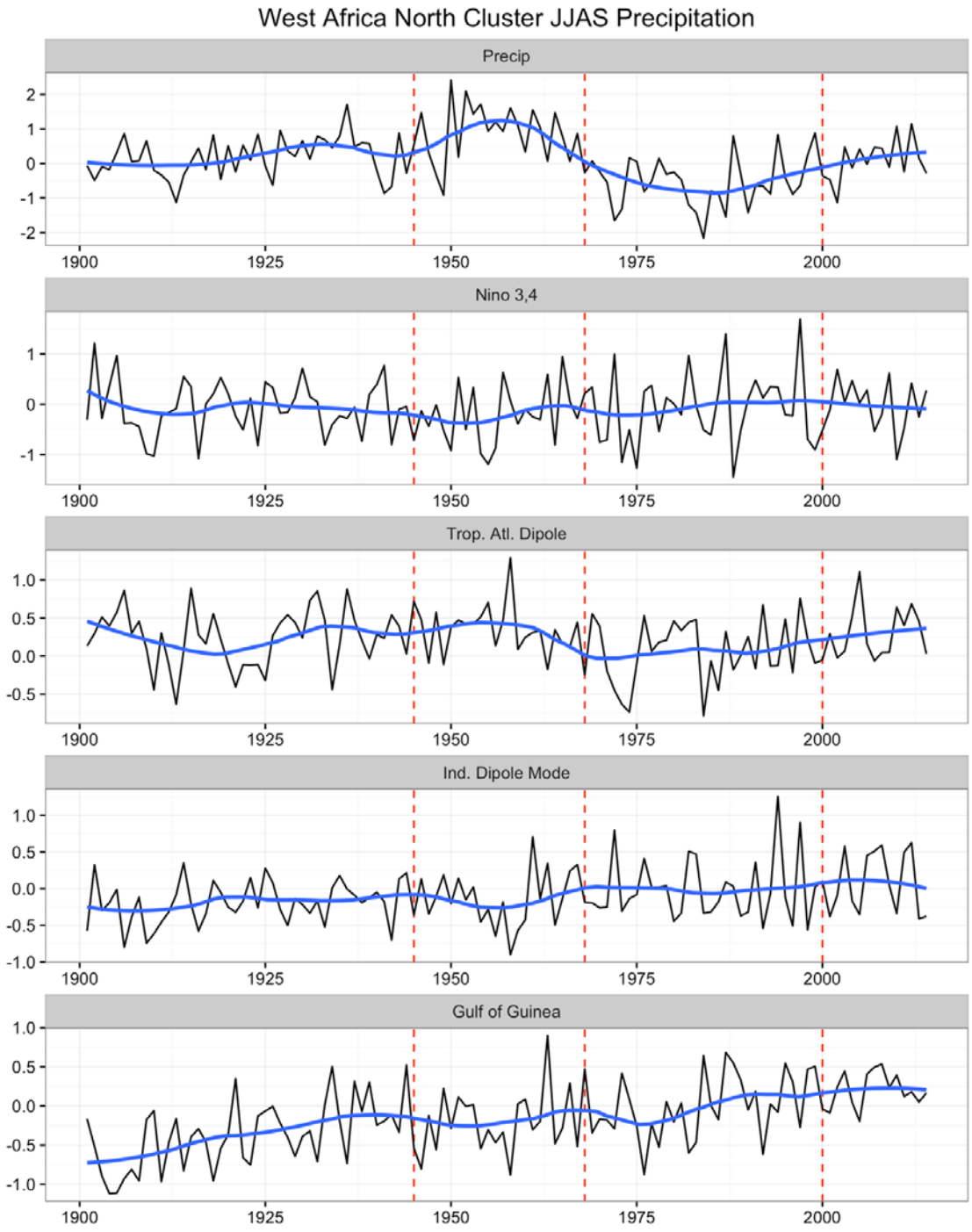

Figure 2-6 Standardized precipitation and SST index (black line) and smoothed values (blue line) for north cluster precipitation (top panel), Nino 3,4 (2nd panel), TAI (3rd panel), DMI (4th panel), and GGI (bottom panel). Red lines show the boundaries of the defined epochs: wet period (1945-1968), drought (1969-2000), and modern period (2001-2014).

Timeseries of the indices and standardized precipitation are shown in Figure 2-6. The top panel shows an anomalous wet period in Sahel during the middle part of the 20th century starting in the early to mid-1940s and lasting until the late 1960s, followed by the Sahel drought lasting until the end of the century. The Nino 3,4 shows the well-reviewed multidecadal pattern of cooler during the middle part of the 20th century and warmer subsequently with a cooling in the early 21st century. The TAI also shows a multi-decadal pattern, and interestingly has been increasing from the 1990s to the present. The GGI shows a similar increasing trend since 1975. The IOD-DMI has a slight positive trend from the 1990s to the present along with increasing variability, with larger and more frequent year-to-year changes. The epochal shifts in precipitation correspond to shifts in the strengths of the SST teleconnections as seen in the BDLM regression coefficients (Figure 2-5**)** and also the temporal variability of the indices in Figure 2-6.

2.4 EPOCHAL COMPOSITES

For the epochal composites, we have lowered the threshold for identifying low and high precipitation years to 0.4 for wet years and -0.4 for dry years in order to identify low years during the mid-century wet period and high years during the Sahel drought. The composites of SSTs for the three epochs and for the season and sub-season periods are shown in the following figures. During the mid-century wet period, the tropical Pacific is anomalously cooler for wet years (Figure 2-7) during all parts of the season, while the dry years show much weaker warming compared to the composites for the entire period (Figure 2-3).

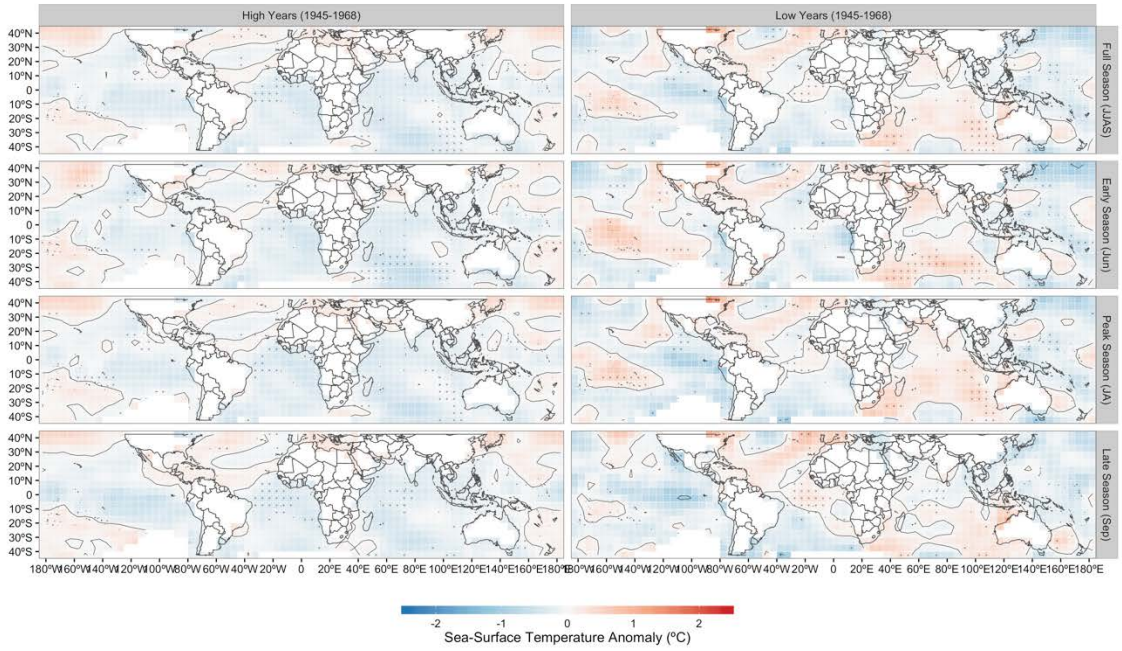

Figure 2-7 Same as Figure 2-3, but for the wet period (1945-1968)

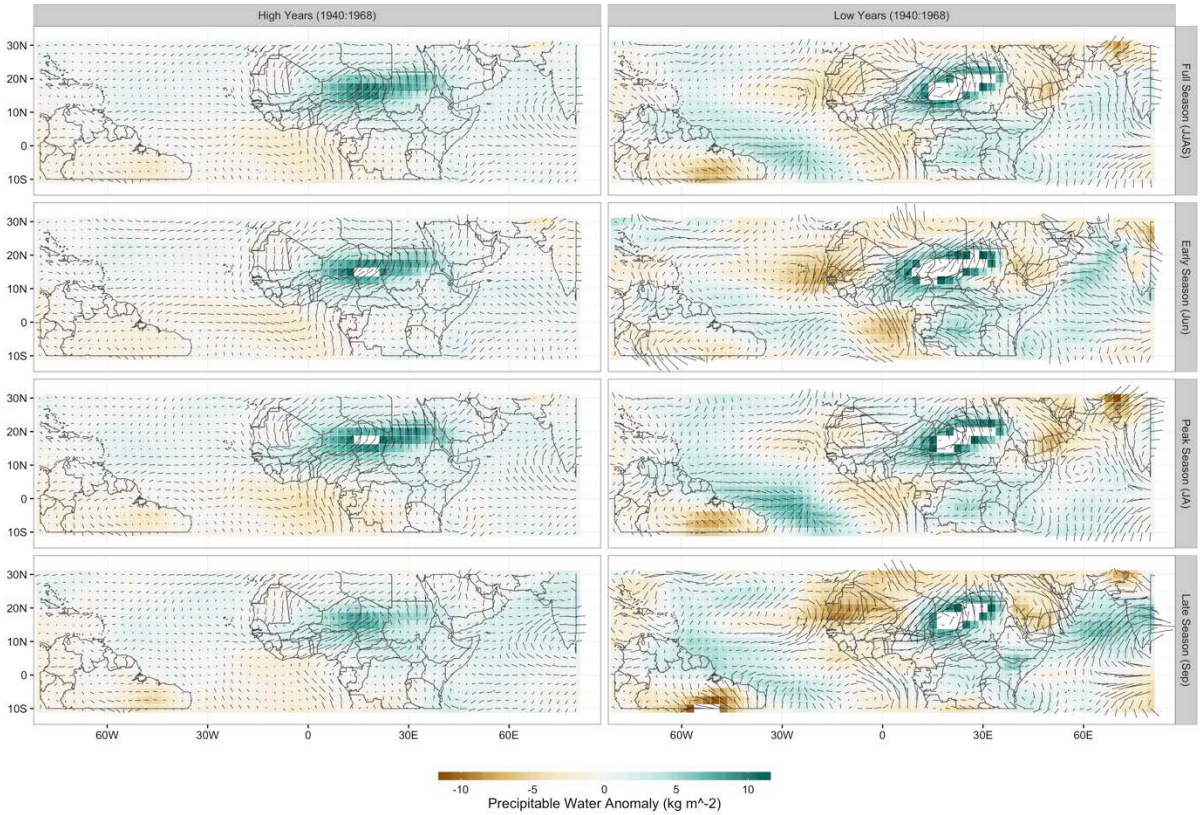

Figure 2-8 Same as Figure 4, but for the wet period (1945-1968)

The Atlantic dipole is not as evident, and the Indian Ocean cool/wet, warm/dry correspondence is also absent compared to the full period analysis (Figure 2-3). This is consistent with BDLM analysis which showed a weakening in the ENSO relationship, a weak Atlantic dipole (Figure 2-5). The full season and September late season also show anomalous behavior in the western Gulf of Guinea, with cooler conditions during wet years, and warmer conditions during dry years. The same anomalous precipitable water and onshore low-level winds seen in the full record are seen in high years, and even in low years (Figure 2-9). Upper level wind anomalies exhibit the same behavior, with anomalies in high years corresponding to the full period, and these same easterly wind anomalies still present in low years.

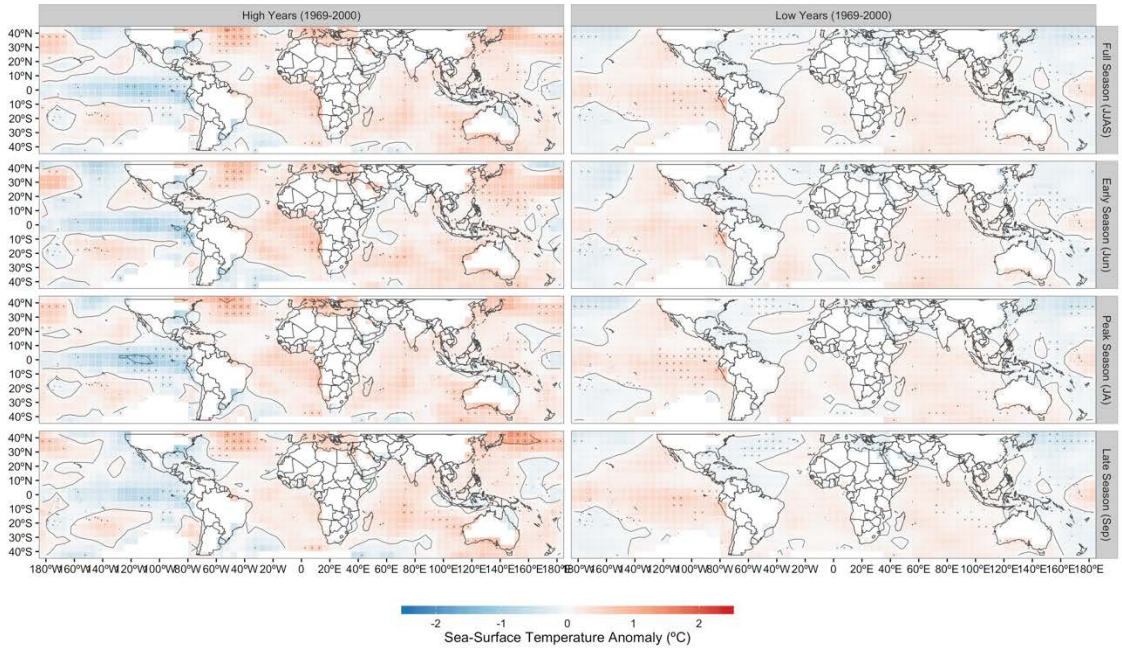

Figure 2-9 Same as Figure 2-3, but for the drought period (1969-2000)

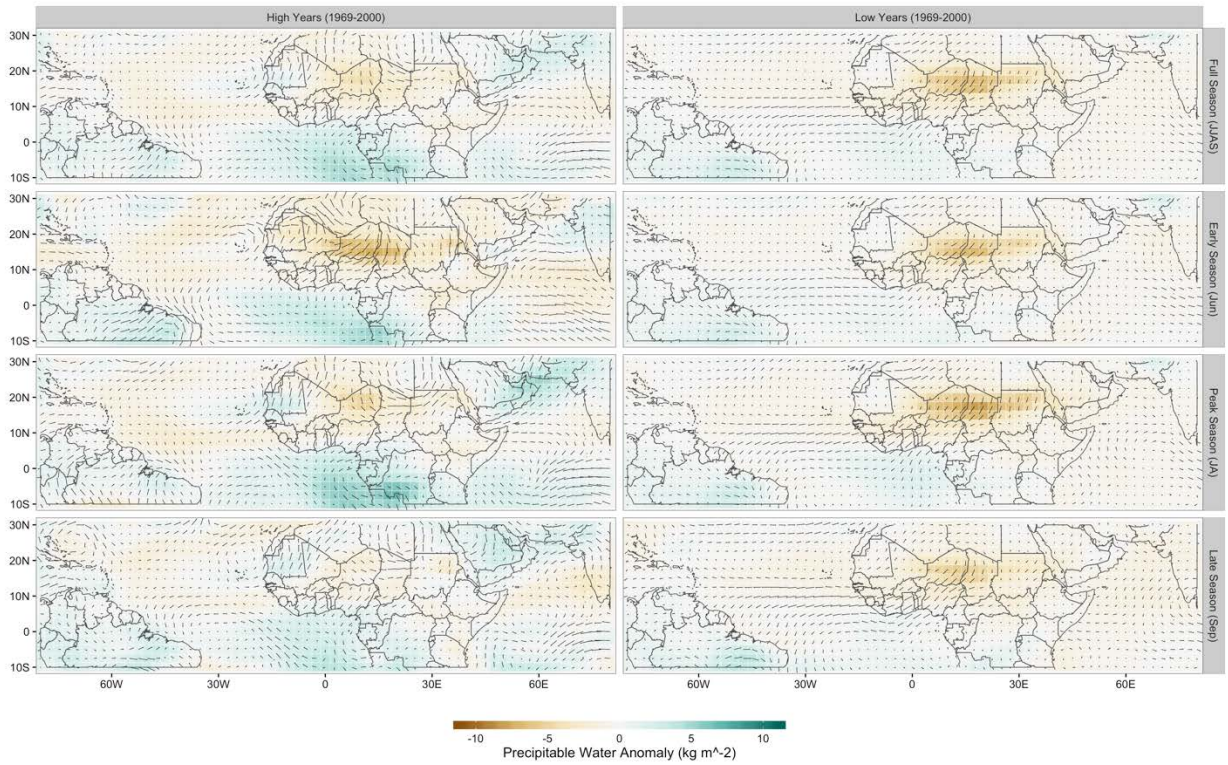

Figure 2-10 Same as Figure 4, but for the drought period (1969-2000)

The SST composites during the Sahel drought epoch (Figure 2-10**)** shows the return of an ENSO pattern in the Tropical Pacific with a cooler eastern Pacific during wet years and a warmer eastern Pacific during dry years. While the Atlantic dipole is not strong, warmer conditions are observed in the Gulf of Guinea during wet years and cooler (though still positive anomalies but weaker) conditions are observed in dry years. The positive precipitable water anomaly present in the full record high precipitation year composites is replaced by a negative precipitation anomaly, strongest during the June early period. In addition, there are also anomalous low-level winds from the north and east extending to the Guinea and Atlantic coasts. The same negative precipitable water anomaly and low-level wind anomalies are seen in dry years, with the precipitable water anomaly persisting through the peak and late periods (Figure 2-11). Upper level wind anomalies during dry years are

very similar to the full record anomalies, but during dry years, the positive easterly anomalies are replaced by weak westerly anomalies.

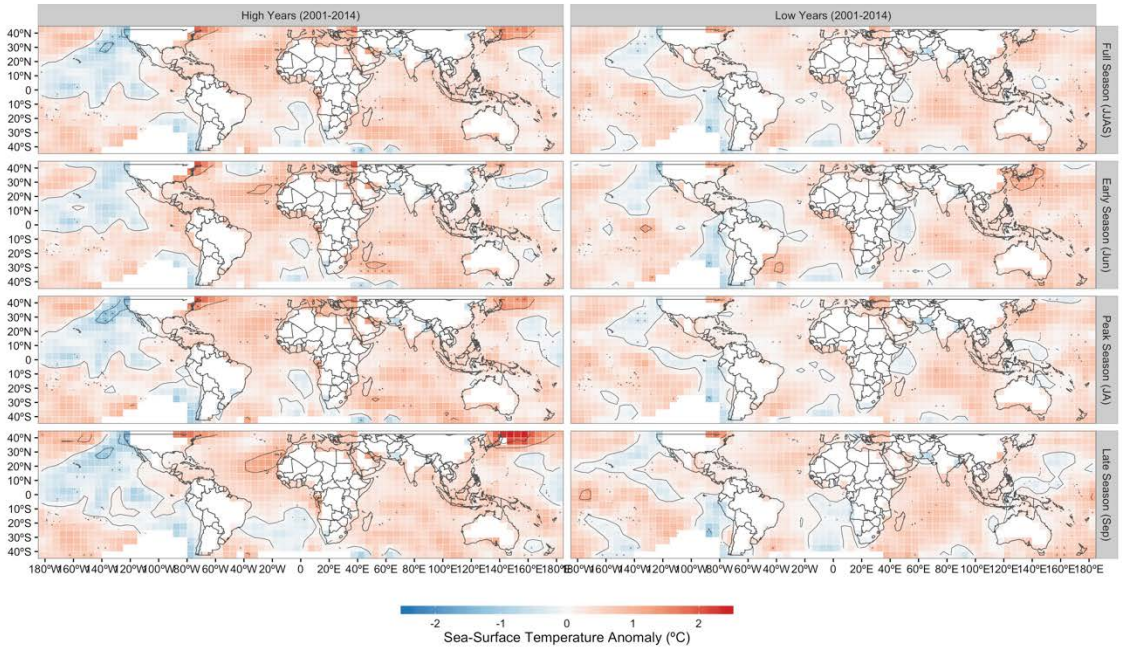

Figure 2-11 Same as Figure 3 but for the modern period (2000-2014).

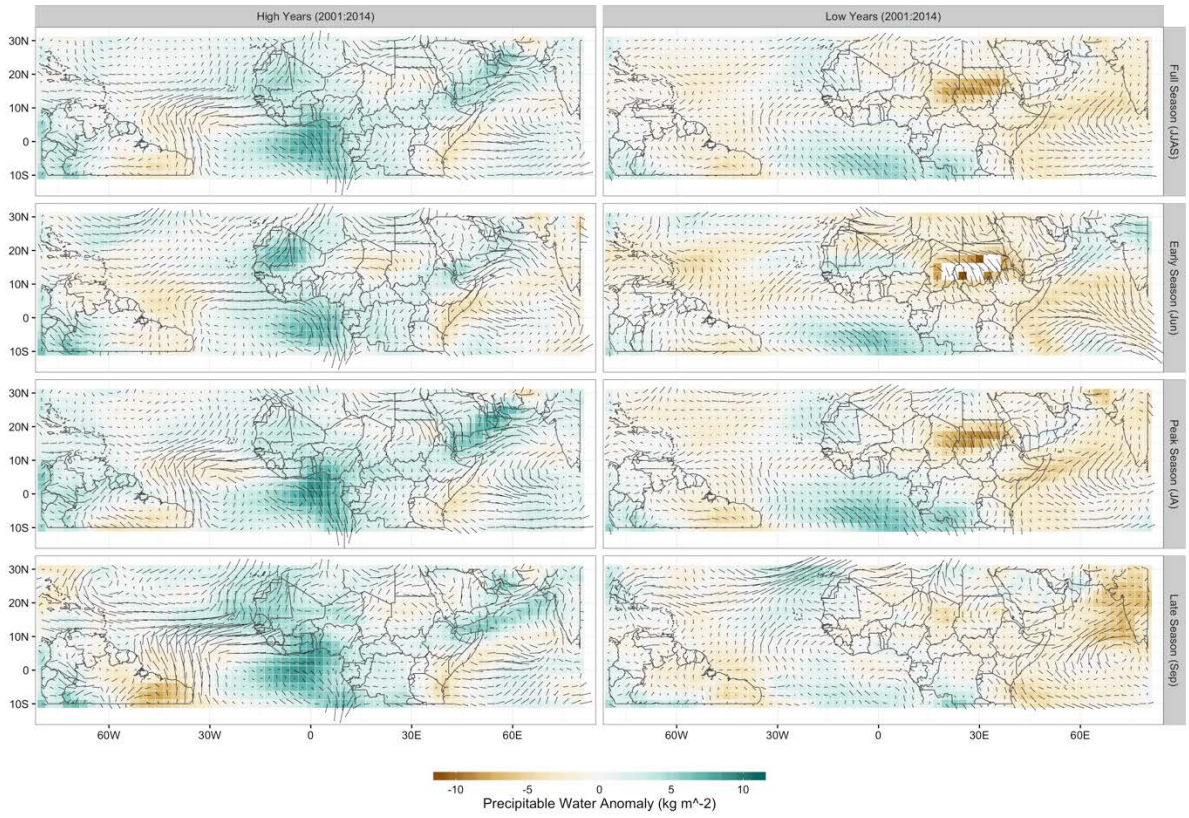

Figure 2-12 Same as Figure 4, but for the modern period (2000-2014)

The SST anomaly composites for the modern period (Figure 2-12) show an increase in temperatures throughout, for both wet and dry years. A consequence of this increase in warming has been the increase in strength of more local SST regions, namely the Gulf of Guinea, and a decrease in the strength of ENSO as seen in the BDLM results. Also, the TAI is increasing to the levels seen in the pluvial wet period, but the precipitation is lower than this earlier wet period. This suggests that the warming in Indian Ocean and Gulf of Guinea could be acting negatively on the precipitation. This has been seen with Indian monsoon precipitation, where La Nina like conditions in the Pacific that historically produced heavy rains are barely producing above normal rains in recent decades (with 2016 as a prime example) (Krishnaswamy et al. 2015; Gill et al. 2015a). Warming in Indian Ocean as a result of anthropogenic planetary wide warming inhibits the strengthening of the circulation that is important for moisture convergence.

2.5 SUMMARY AND DISCUSSION

In this paper, we performed a systematic analysis of 112-year-long (1901-2012) and investigated the variability of Sahelian summer monsoon precipitation and the strength of its teleconnections at multidecadal, seasonal and subseasonal time scales. The wet / dry precipitation anomalies extend beyond the Sahel in to northern East Africa, and over India, which suggests coherent large-scale variability of precipitation influenced by large-scale teleconnections from SST forcings. The coherence is strong with East Africa during the peak season and in the late season with India. Similar relationships are seen from composites of anomalous precipitation during La Nina and El Nino years. As with the Indian monsoon, wet years strongly correspond with La Nina conditions in the tropical Pacific and dry years with in El Nino conditions, albeit weakly. Through the BDLM we showed that there is multi-decadal variability in the relationship between precipitation and large-scale SST teleconnections. Over the last century the relative strengths of the tropical Pacific, the Indian Ocean, the tropical Atlantic, and the Gulf of Guinea have varied, with changes corresponding to epochal shifts in precipitation. During the mid-century wet period, Nino 3,4, the TAI, and the IOD-DMI are all relatively weak, with the GGI exhibiting a negative relationship. The subsequent Sahel drought sees a continued weak TAI and IOD-DMI, with Nino 3,4 strengthening and the GGI weakening before going positive. In the modern period, Nino 3,4 has again weakened, the TAI and IOD-DMI have returned to the average strength values, and the GGI has continued to increase in strength in the positive direction. This show that during epochal shifts, there are changes in the relative strengths of multiple SST teleconnections rather than a change in a single dominant SST teleconnection. Circulation field anomalies for the full record clearly showed the large-scale dynamics of precipitation in this region, namely low-level moisture transport inland from the Atlantic and Gulf of Guinea, and the strength and location of the upper-level TEJ. For both the wet period and dry years these features correspond as expected and persist in the opposite years. Onshore low-level wind anomalies and easterly upper level wind anomalies are also present in low years during the wet

period while northeasterly low-level wind anomalies and westerly upper level winds are present in high years during the drought. Increases in global SSTs in modern times has led to a new epoch in summer monsoon precipitation, where the relationships with the Gulf of Guinea and Indian Ocean have strengthened and the relationships with ENSO and the Tropical Atlantic have weakened.

This study offers insights into the rich variability of Sahel rainfall at multiple time scales. The asymmetry of ENSO teleconnections during wet and dry years subseasonal is quite interesting. Similarly, the temporal variability of the teleconnections. Also interesting is the similarity of variability and teleconnections exhibited by the Indian monsoon rainfall. All of these argue for a coherent space-time variation of rainfall over Africa and India that is hitherto not fully recognized. The insights uncovered from this study, offer exciting prospects to improve the skill in seasonal forecasting, especially to develop statistical models conditioned on the state of the large-scale SST drivers. These can be used in statistical post-processing of dynamical forecasts. Furthermore, these can be used to develop robust multidecadal projection precipitation using wavelet and Hidden Markov Models (HMM) e.g., (Erkyihun et al., 2016; Erkyihun et al., 2017) – wherein the large-scale SST indices are modeled as wavelets of HMM and the precipitation simulated conditionally. In addition several interesting scientific questions for inquiry emerge from this study – How does the strength of teleconnections vary in a warmer climate?; Are the variations in SST drivers related? If so how? Are there space-time variability of precipitation in the Africa – India region part of a global tropics-wide precipitation variation? Probing climate model simulations and idealized model experiments can help provide insights into these questions.

3 FORECASTING AND EVALUATING LIVESTOCK HEAT STRESS USING A STOCHASTIC WEATHER GENERATOR

3.1 INTRODUCTION

Few studies have investigated the impact of heat stress on livestock within smallholder systems, which directly support 600 million people in the developing world (Thornton et al. 2006). These studies have, at least in part, been limited by a lack of accurate and reliable data on both environmental conditions and livestock health. Heat stress, the physiological response to direct environmental exposure, manifests as an increase in an animal's core temperature, leading to decreases in metabolism, production, and in severe cases, death. Heat stress burden and risk have been historically determined using index-based measures, with empirically derived thresholds for severity. The most widely adopted measure is the thermal-humidity index (THI) (Thom 1959; NOAA 1976). The advantages of the THI are that it requires only temperature and humidity as inputs, two commonly measured climate variables, is easily calculated, and has risk levels that can easily be communicated to practitioners. In proposing replacements, studies have highlighted the shortcomings of the THI: it fails to account for the influence of additional climate variables including winds, solar radiation, and access to shade, it is a point measure of heat stress in time and does not account for cumulative effects, and the thresholds required to determine impacts are dependent on animal breed and physical characteristics (Gaughan 2006; Hahn et al. 2009; Mader et al. 2010; Thompson et al. 2013). Two of the approaches that have been taken to address these shortcomings are: alternative index-based measures that incorporate additional climate variables, and energy-balance approaches that explicitly model the animal's heat budget. Index-based measures include the Heat Load Index (HLI) (Gaughan et al. 2008), Comprehensive Climate Index (CCI) (Mader et al. 2010), and the Livestock Weather Safety Index (LWSI) (Eigenberg et al. 2005; Brown-Brandl et al. 2006; Eigenberg et al. 2007). These three measures provide point estimates of heat stress. To account for the cumulative effects of heat stress, the Accumulated Heat

Load (AHL) index uses the HLI with fixed rates of increase or decrease when above or below critical thresholds (Gaughan et al. 2008; Mader et al. 2010). Energy balance approaches explicitly model the cattle heat budget, building on conceptual models laid out in Finch 1986 and Monteith and Unsworth 2014. These energy balance models take in external inputs including environmental temperature, winds, and solar radiation (direct exposure to the sun), and combine them with heat cattle produce internally through digestion and reproduction. This approach addresses several issues found with using an index-based measure, implicitly incorporating cumulative effects, and removing the need for breed and location-based, empirically derived thresholds. It does however introduce the need for accurate physiological cattle data. Either approach requires complete weather sequences as forcings, which can be difficult to obtain for many study regions. Stochastic weather generators (SWGs) present an attractive method for producing a rich variety of forcing data. Additionally, the inclusion of covariates in SWGs provide a flexible tool for generating weather sequences for use in short-term and seasonal projections conditioned on seasonal climate forecasts, for use in livestock planning and management decisions.

In the developing world, often in locations where livestock are a major livelihood, research into livestock heat stress burden has been limited by a lack of weather data. Although accurate physiological cattle data are needed to define critical thresholds for use with the index-based methods, or for inclusion in heat-budget models, quantifying the variability of historical and future heat stress burden from variability in weather can still provide useful information about potential impacts using current best-estimate values. The use of a SWG provides a means of quantifying this variability and providing an initial assessment of heat stress burden in these regions. SWGs also provide a means of obtaining the climate variables required for more calculating more complex measures of heat stress which is potentially more reflective of actual on-the-ground conditions.

While weather generators have found acceptance in the agricultural community, they have primarily been used to examine crop production. Applications

to livestock related issues have been limited to one study, Parsons et al. 2001b,a, who used stochastic weather output to examine the impact of changing climate on livestock production across Britain using several process models. Our research here marks the first application of SWGs for livestock health. In this paper, we demonstrate the utility of a SWG in quantifying livestock heat stress risk in data-scarce regions. We highlight the ability to incorporate in a wide variety of covariates which can be used to produce current-condition estimates of heat stress, subseasonal to seasonal forecasts, and future climate projections. This flexibility provides a robust tool that can offer actionable information to both decision makers and planners.

The chapter is organized as follows. A review of SWGs is first presented, followed by a description of the data and study region. The SWG model used is next presented, followed by results from conditional and unconditional simulations, with a discussion and future extensions to conclude the paper.

3.2 STOCHASTIC WEATHER GENERATOR OVERVIEW

Stochastic Weather Generators produce synthetic weather sequences that are statistically consistent with the historical record. The original single-site weather generator of Richardson 1981 and Richardson and Wright 1984 model precipitation occurrence as a chain-dependent process (Katz 1977) and amount using probability distributions. This two-step approach allows the generator to capture the length and frequency and magnitude of wet and dry spells. Temperature is modeled using linear time series models. This approach well captures historical climatology, but is poor at capturing extreme events. Incorporating in appropriate covariates to condition the weather generator runs in a particular season offers a way of improving a weather generator's performance. Examples of appropriate covariates include ENSO phase (Trenberth 1997) and other climate teleconnections (Grondona et al. 2000; Andrés Ferreyra et al. 2001; Katz 2002; Wilby et al. 2002; Meza 2005). Extensions to this approach have provided methods for perturbing parameters,

applying seasonal correction factors, adjusting input data, and spectral bias corrections (Hansen and Mavromatis 2001; Caron et al. 2008; Kilsby et al. 2007; Qian et al. 2010). In addition, weather generator parameters can be interpolated in space using locally-weighted regression to produce simulations of weather at arbitrary locations (Wilks 2008). The development of non-parametric weather generators has offered improvements in capturing nonlinearities between variables and sites. Non-parametric approaches include k-nearest neighbor (k-NN) bootstrap resampling (Brandsma and Buishand 1998; Rajagopalan and Lall 1999; Buishand and Brandsma 2001; Beersma and Adri Buishand 2003; Yates 2003; Sharif and Burn 2007) and kernel density based estimators (Rajagopalan et al. 1997; Harrold 2003). Further modifications have been made to improve performance. Initially applying a clustering algorithm to identify regions of similar climatology (Caraway et al. 2014) has improved simulations in complex terrain. Using a semi-parametric approach with Markov chain modeling has improved the ability to capture the duration of wet and dry spells, and including the ability to incorporate in seasonal forecasts (Apipattanavis et al. 2010) or multi-decadal projections (Podestá et al. 2009) enables weather generators to be used in forecasting applications. Despite these developments, the primary limitations of this class of weather generators is that it is difficult to produce values outside the range of historically observed weather and difficult to produce simulations at locations lacking historically observed weather.

Stochastic weather generators using generalized linear models (GLM) offer several advantages over traditional approaches. GLMs are a variation on traditional linear regression models which allow an appropriate model to be used for each climate variable, for example normal distribution for temperature or gamma distribution for precipitation amount which cannot be less than zero. Each model is constructed by specifying the appropriate link functions. This form of weather generator was first demonstrated by Stern and Coe 1984 and improved by Yang et al. 2005. The GLM structure is flexible in that it can incorporate the appropriate set of covariates, such as short-term or seasonal forecasts, or multi-

decadal projections. In Kim et al. 2012 seasonal precipitation was used as a covariate to generate weather sequences at a single site and in Verdin et al. 2014 a space-time GLM weather generator was used to generate daily weather ensembles over a sub-basin in Argentina, consistent with seasonal climate forecasts and future climate projections. The utility of this approach was demonstrated by applying it to agricultural management in the Argentinian Pampas.

Numerical weather prediction models, seasonal climate forecasts, and future climate projections can all be used to obtain these covariates, leveraging their forecast skill and the ability to produce many realizations. Projections of future climate from global climate models (GCMs) are mostly at coarse spatial scales, while resource management decisions require information at the local scale. Covariates from GCMs at the coarse spatial scale can be used with SWGs to generate ensembles of daily weather at the fine scales required for decision making (e.g. Verdin et al. 2018). These can be used to drive process models, such as agricultural or hydrologic models to relate daily weather to measures relevant for local decision making. Thus, SWGs have been used as a downscaling approach. In particular, for livestock management, stochastic weather ensembles can provide seasonal projections of heat stress which will enable efficient decision making by livestock owners, including purchase and sale of cattle, fodder, investment in shelters, and identify locations with the best conditions for seasonal migrations. They can also serve as the basis for insurance programs which can offer herders and farmers a financial buffer against uncertain future conditions and climate change. These index and climate based insurance programs are becoming increasingly popular and effective tools in addressing the impacts of climate in the developing world (Greatrex et al. 2015; Surminski et al. 2016).

3.3 STUDY REGION

The focus for this study is the Sahelo-Sudanian region of West Africa (Figure 4-1), where agricultural production in aggregate supports 60% of the active

population and is responsible for 40% of the gross national product (Battisti and Naylor 2009).

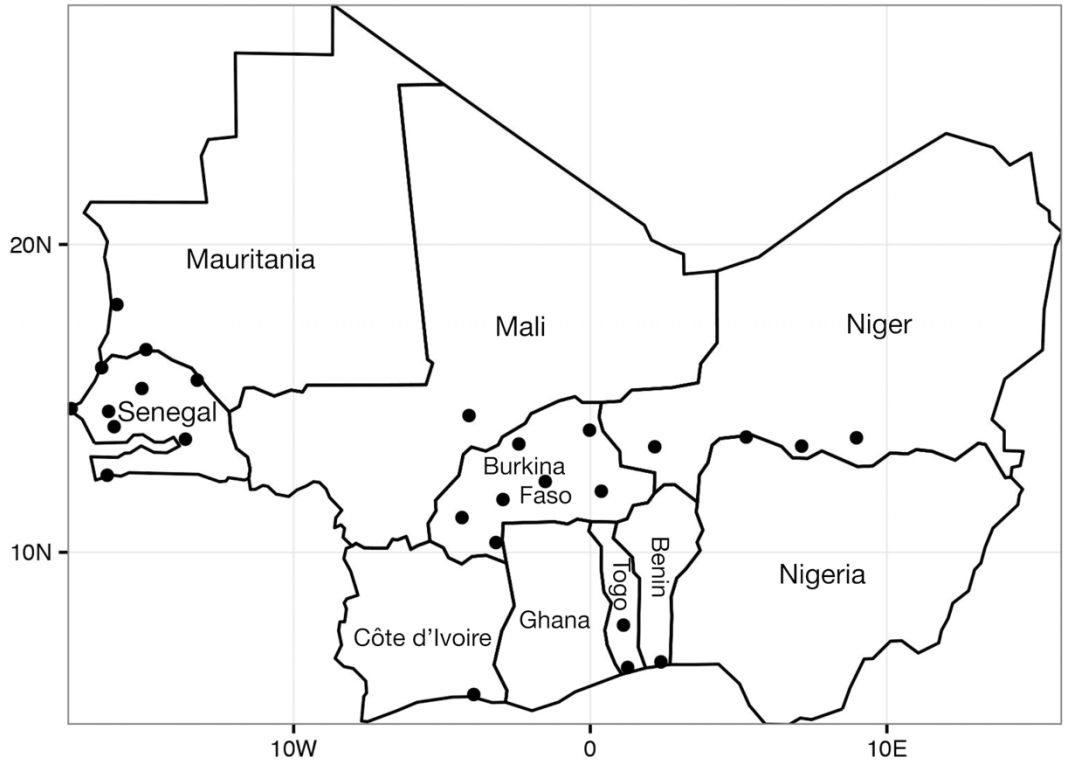

Figure 4-1 Study region showing the 26 weather observations stations used (black dots).

Quantifying heat stress burden and livestock risk in the developing world, especially in this region, has been limited, with little understanding of the susceptibility of native breeds to heat stress. This lack of knowledge has been highlighted by the International Livestock Research Institute (ILRI), who states "the impacts of increased frequencies of extreme heat stress on existing livestock breeds are not known, nor do we know if there are critical thresholds in the relationship between heat stress and physiological impacts." Easterling and Apps 2005 state: "Confidence in the ability of livestock producers to adapt their herds to the physiological stresses of climate change is difficult to judge. The absence of physiologically based animal models with well-developed climate components

suggests a major methodological void." Additional complexity is introduced when considering a broad region such as sub-Saharan Africa where there is a large heterogeneity in livestock systems in terms of breeds, methodology, and adaptive capacity. Adaptation and selective breeding have produced cattle breeds with greater resistance to local conditions including climate and disease. The resilience of these breeds to heat stress has not been well quantified, with the threshold values commonly used with heat stress indices developed for intensively managed cattle in the developed world. In addition, there has been limited effort in developing accurate energy budget models for these breeds. Application of a SWG to livestock heat stress in this region can offer baseline heat stress burden, and the ability to forecast heat stress burden at multiple time scales. In this paper, we demonstrate the utility of a SWG in quantifying livestock heat stress risk in data scarce regions. We highlight the ability to incorporate in a wide variety of covariates which can be used to produce current-condition estimates of heat stress, subseasonal to seasonal forecasts, and future climate projections. This flexibility provides a robust tool that can offer actionable information to both decision makers and planners.

3.4 DATA

Daily weather data were obtained from the HadISD dataset, a quality-controlled version of the NOAA Integrated Surface Dataset (ISD) from the UK Met Office. These data are collected from over 100 sources and provide 3-hourly observations at each location (Dunn et al. 2012). Daily minimum, maximum, and mean values were calculated for temperature, relative humidity, and wind speed along with daily precipitation totals. Stations were filtered by the percent of missing data in the study period 1973-2012 and those with at least 90% coverage were retained. Stations with gaps in the data longer than a month were removed from this subset. The 26 stations meeting these criteria are shown in (Figure 4-1). The NCEP-NCAR reanalysis (Kalnay et al. 1996) provided 3-hourly surface solar

radiation data by selecting the values from the closest grid box for each observation station location. Daily mean solar radiation was calculated from these data.

3.5 MODELING FRAMEWORK

3.5.1 HEAT STRESS MEASURES

To demonstrate the utility of a SWG in quantifying heat stress burden, two of the heat stress measures, THI and HLI, are used to quantify livestock heat stress burden. While the limitations of THI are understood, using only temperature and humidity, it remains a widely used heat stress measure. The HLI, which includes winds and solar radiation in addition to temperature and humidity, shows the ability of a SWG to accurately model these additional climate variables. The SWG used in this paper produces the climate variables required to calculate other index-based heat stress measures and energy-balance based measures as well, but for clarity we chose to present results for just THI and HLI.

$$\mathrm{THI} = 0.8\overline{\mathrm{T}_a} + \frac{RH}{100} \times (\overline{\mathrm{T}_a} - 14.4) + 46.4$$

HLI is more robust as it includes solar radiation and wind speed in addition to temperature and humidity. It is calculated using:

$$\mathrm{B_gT} = \left(1.33\mathrm{T}_a - 2.65\sqrt{\mathrm{T}_a} + 3.21 \log S_r + 1\right) + 3.5$$

$$\mathrm{S.\,B_gT} = 1/\left(1.33\mathrm{T}_a + e^{\frac{-\mathrm{B_gT}-25}{2.25}}\right) + 3.5$$

$$\mathrm{HLI.\,lo} = 1.3\mathrm{B_gT} + 0.28(\mathrm{RH} - \mathrm{WS}) + 10.66$$

$$\mathrm{HLI.\,hi} = 1.55\mathrm{B_gT} + 0.38(\mathrm{RH} - 0.5\mathrm{WS}) + e^{(2.5-\mathrm{WS})} + 8.62$$

$$\mathrm{HLI} = \mathrm{S.\,B_gT} \times \mathrm{HLI.\,hi} + \left(1 - \mathrm{S.\,B_gT}\right) \times \mathrm{HLI.\,lo}$$

where BgT is a calculated black globe temperature. S.BgT is a blending function used with HLI.lo and HLI.hi to ensure a smooth function.

The seasonal climate of the region is largely controlled by the West African Monsoon (WAM), a low-level southwesterly flow existing during the boreal summer. This flow advects moisture onshore, and the strength of the WAM impacts both the magnitude and location of summer rainfall. The WAM also affects seasonal humidity and temperature. This can be seen during monsoon onset and retreat (in the spring and fall respectively), with changes in the timing of these transition periods affecting local climate conditions directly, and livestock heat stress indirectly.

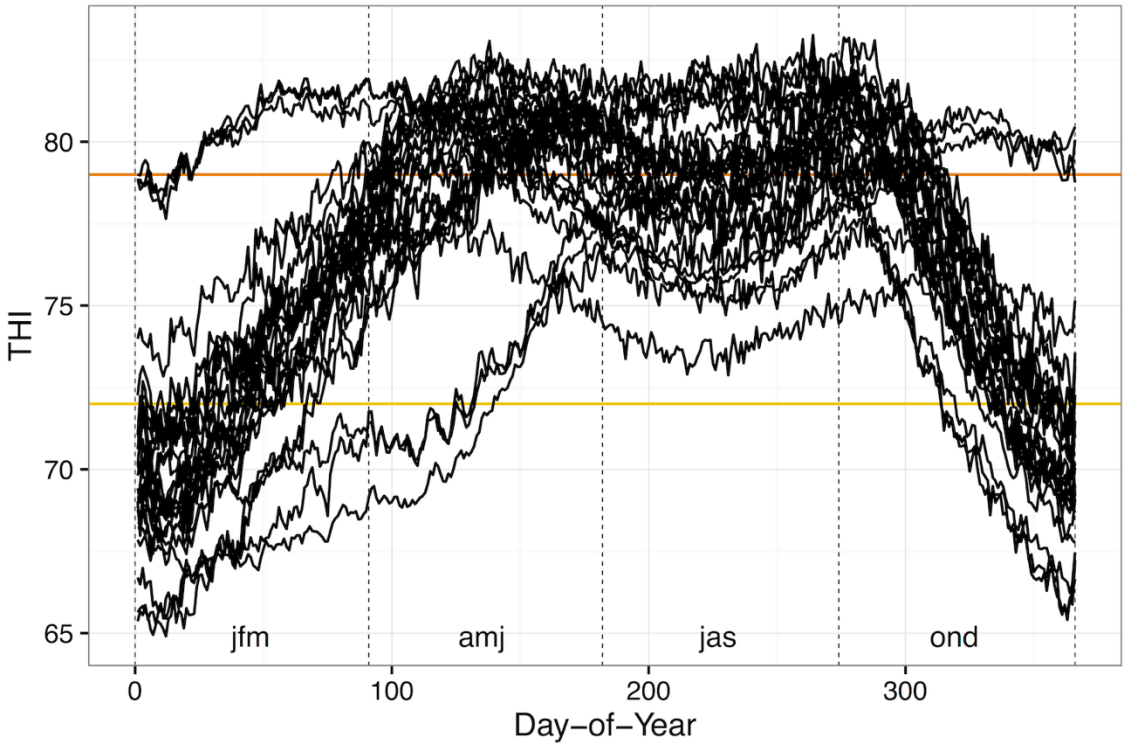

Figure 4-2. Median climatology for individual stations in the study region. Line indicate commonly-used thresholds for mild (yellow) and moderate (orange) heat stress.

From the climatology of THI at all weather observation stations (Figure 4-2) the values peak in the spring, with a dip in values through the summer, and a second peak in early fall - monsoon rainfall is responsible for this dip in summer values. Timing of monsoon onset and retreat can shift the spring and fall peaks earlier or later in the season. In addition to THI (or other heat stress measure) magnitude, additional insights into heat stress related risk can be provided by the number of exceedances above critical thresholds, and the duration of exceedance, or number of consecutive days above the threshold provide additional insights into the heat stress related risk. These three measures of heat stress - magnitude, exceedance counts, and consecutive days of exceedance, can be used to quantify the severity of heat stress. Ensembles of daily weather sequences can be used to provides a rich variety of heat stress measures to capture the risk to livestock health. This understanding of the severity of heat stress can serve the smallholder livestock community at large, and especially the pastoralists who migrate throughout the year. Current decision making about when and where to migrate are based on past experience, and available reports about pastureland conditions and water access. Associated information about heat stress could provide an additional factor to consider in this decision making. There have also been efforts to develop index-based insurance programs that rely on local conditions to assess risk and issue payouts when conditions reach a certain threshold. Measures of heat stress could also be incorporated into these programs to more accurately represent conditions on the ground (Greatrex et al. 2015).

3.5.2 STOCHASTIC WEATHER GENERATOR

A stochastic weather generator as described in Furrer and Katz (2007, 2008) and implemented in Kim et al. 2012; extended to space-time generation of precipitation and temperature (Kleiber et al. 2012, 2013), and integrated into a robust space-time weather generator in (Verdin et al. 2014) is employed to investigate the statistical variability of and to model minimum and maximum

temperature, minimum and maximum relative humidity, minimum and maximum wind speed, precipitation occurrence and intensity, and mean solar radiation, climate variables required to calculate heat stress measures. In this research, extension of the SWG to relative humidity and winds is an important contribution. Each climate variable is modeled using a Generalized Linear Model (GLM) (McCullagh and Nelder 1989) with an appropriate 'link function'. The GLM framework provides a method of linking the response variable to a set of covariates. The response variable is assumed to be a realization from any distribution in the exponential family with a set of parameters. A smooth and invertible link function transforms the conditional expectation of $\boldsymbol{Y}$ to the set of covariates, $\boldsymbol{X}$ with model parameters $\boldsymbol{\beta}$ and variance of error ε, which is assumed to be normally distributed with mean zero and variance $\boldsymbol{\sigma^2}$.

$$G(E(Y)) = \eta = f(\mathbf{X}) + \varepsilon = \mathbf{X}\boldsymbol{\beta}^{\mathbf{T}} + \varepsilon$$

An iterated weighted least squares method is employed to estimate the model parameters, $\boldsymbol{\beta}$. An appropriate link function, G(.), is selected based on the assumed distribution of Y. All statistical analysis and modeling is performed using R (R Foundation for Statistical Computing). Details for each model can be found in Appendix A. The structure of these models for precipitation occurrence, precipitation intensity, and temperature follow on past SWGs. Relative humidity, wind speed, and solar radiation were added to calculate heat stress measures, but are also applicable to other measures in agriculture and public health. With relative humidity, logistic regression approach was taken, using a logit link function in the GLM to ensure values between 0 and 100. Day-of year variance for all variables is modeled using harmonic terms to capture the changing seasonal variance following the form:

$$\ln(V_{doy}) = \mu + \beta_{i,1} C_t + \beta_{i,2} S_t$$

where $\ln(V_{doy})$ is the log of the day-of-year mean variance and C_t and S_t are the harmonic terms as before. When simulating, the error is sampled from Normal distribution with the estimated variance from the above model on the day-of-year of interest. This addition greatly improves the weather generator's ability to capture processes with a seasonally changing variance. The structure of these models with the addition of covariates can also be found in Appendix A. Seasonal covariates: mean temperature, mean relative humidity, and total precipitation can be obtained using forecasts providing a way of generating daily weather for an upcoming season. The International Research Institute for Climate and Society (IRI) at Columbia University produces seasonal tercile forecasts of precipitation giving the probability of being above average, average, or below average. These probabilities can be used to conditionally sample historical observations and obtain seasonal covariates. For a given year, the probabilities of precipitation above average, average, or below average precipitation are obtained from IRI seasonal forecasts. Historical seasonal precipitation totals are also classified as above average, average, or below average, and the desired number of years are sampled using the probabilities as weights, The seasonal temperature, relative humidity, and precipitation from the sampled years are then used in the weather generator as covariates. The daily values for the start of the season for each sampled year are used as initial conditions.

3.6 RESULTS

3.6.1 CONDITIONAL SIMULATION

An ANOVA table showing the statistical significance for all model coefficients can be found in Appendix B. It can be seen that all the model covariates are statistically significant at 90% confidence level and higher. We generated daily weather sequences at Ouagadougou, Burkina Faso, conditioned with contemporaneous covariates: monthly or seasonal total precipitation, mean, minimum, and maximum temperature, and mean, minimum, and maximum

relative humidity, which we call, conditioned simulation, and for comparison we generated from a model without the covariates, which we call, unconditioned simulation.

3.6.2 HISTORICAL SIMULATION

We simulated daily weather in unconditional and conditional modes. Each simulation was of the same length as the historical data (i.e., 40 years) and 200 simulations were made. For the conditional simulations, monthly covariates were used. Basic monthly statistics of the variables, mean, variance, etc., are generally well captured (Furrer and Katz 2007; Verdin et al. 2014) therefore, here we show the probability density function of the variables from the simulations and compare it with that of the historical data. Figure 4-3 shows the PDF of simulations (in gray) and that of the historical data (solid line) for each season and for the nine variables. The PDFs were computed nonparametrically using kernel density estimators (Silverman 1986).

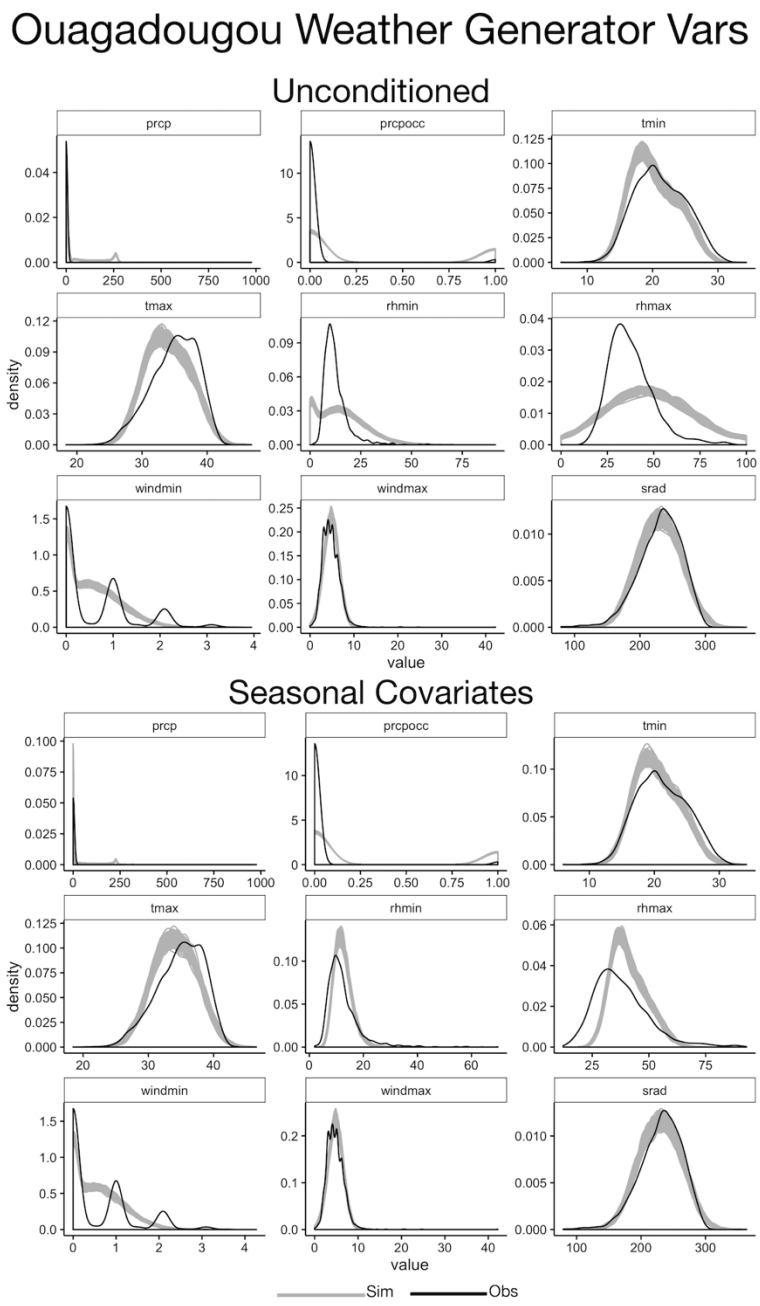

Figure 4-3. Unconditioned and conditioned seasonal weather generator variables for Ouagadougou, Burkina Faso for January - March. From left to right, top to bottom: precipitation intensity (mm/day), precipitation occurrence (0: no rain, 1: rain), minimum temperature (°C), maximum temperature (°C), minimum relative humidity (%), maximum relative humidity (%), minimum windspeed (m/s), maximum windspeed (m/s), and solar radiation (W/m^2/d).

There is an overestimation of the number of rainy days in all cases, with even slightly higher overestimates with conditioning. The conditioned runs do a better job at capturing low-magnitude rain events that are largely missed by the unconditioned runs. Low minimum temperatures are better captured by the conditioned runs, but still fail to capture the peak in temperatures between 23 and 25ºC. Conversely, maximum temperature is overestimated between 32 and 35ºC, with conditioning shifting the distribution to the right towards higher temperatures. There is still an overestimation of lower maximum temperatures and an underestimation of higher maximum temperatures, but conditioning provides some improvement over the unconditioned runs. The greatest impact of conditioning can be seen with relative humidity, in both the minimum and maximum cases. The unconditioned runs are not able to capture the bimodal behavior, while there is marked improvement in the conditioned runs. Wind speed and solar radiation were not directly conditioned for any simulation, but as they are dependent on precipitation, temperature, and relative humidity, there is some improvement in skill imparted when conditioning these variables. The distributions of minimum and maximum wind speed are rough given measurements only at discrete intervals, however, the weather generator is able to capture the underlying smooth distribution quite well. The distribution of solar radiation is also well captured.

3.6.3 HEAT STRESS

The heat stress index THI was computed and the PDFs from the unconditional simulations for the four seasons are shown in Figure 4-4. It can be seen that the simulation PDFs are wider in that they generate lower and higher values of the index.

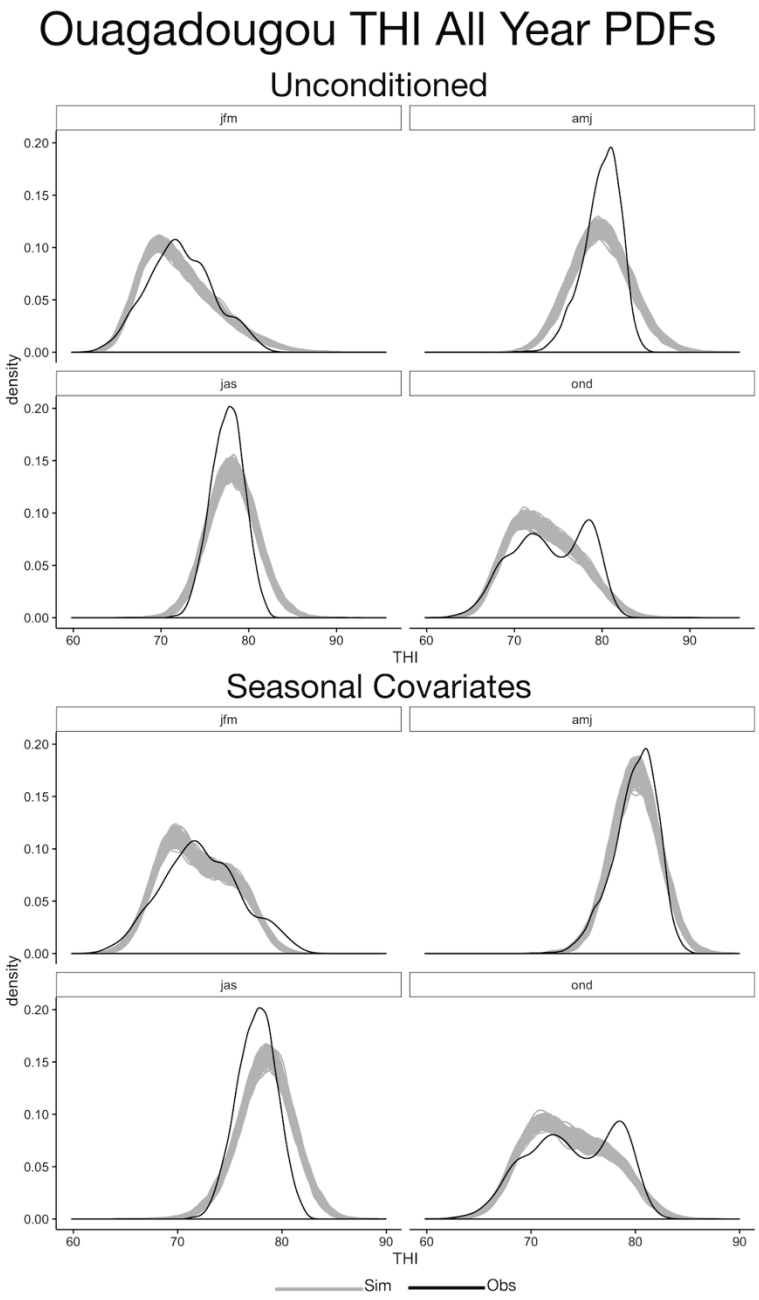

Figure 4-4. Unconditioned and conditioned seasonal thermal-humidity index for Ouagadougou, Burkina Faso. From left to right, top to bottom the panels show the distribution of daily THI for: January - March, April - June, July - September, and October - December. The black line shows the THI calculated from observations, and the gray lines show the THI calculated from weather generator runs.

There is also a tendency for simulating lower THI values as the simulated PDFs are shifted slightly to the left relative to the observed. The distributions of heat stress

indices are well captured with conditioning Figure 4-4 (bottom panel) as compared to the unconditioned (top panel). In January - March (top panel top left and bottom panel top left) conditioning better captures THI values between 75 and 80. While conditioning well captures values at the low end of the distribution, there remains a low bias in the peak as compared with the observed. In addition, while THI values at the high end of the distribution were overestimated without conditioning they are underestimated with conditioning. In April-June (top panel top right and bottom panel top right) conditioning greatly improves the fit, matching the peak seen in the observed, capturing well the low side of the distribution, and reducing the overestimation on the high side of the distribution. June - September (top panel bottom left and bottom panel bottom left) doesn't see much change between the unconditioned and conditioned distributions aside from slight improvements on the low end. Similarly, October - December (top panel bottom right and bottom panel bottom right) show slight improvements with conditioning better capturing the dual peak seen in the observed. The effect of conditioning can best be seen when examining individual years rather than multiple years in aggregate.

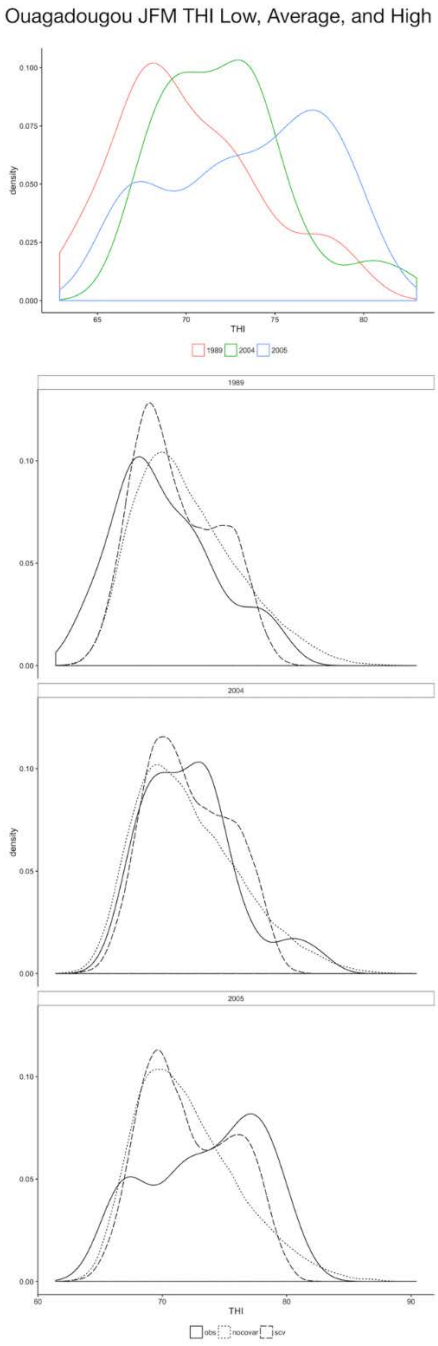

Figure 4-5. January - March THI for low, normal, and high years showing observed, unconditioned, and seasonally conditioned PDFs. The top panel shows the distribution of daily THI from January - March observations for a low year (red), an average year (green), and a high year (blue). The second panel shows the distribution of daily THI from January - March in a low year from observations (black line), unconditioned weather generator runs (small dashed line), and conditioned weather generator runs (large dashed line). The third panel shows the same as the second but for an average year. The fourth panel shows the same as the second, but for a high year.

Figure 4-5 shows the PDFs of THI for 1989 (low year), 2004 (average year), and 2005 (high year) for the January - March season. The top panel shows the distribution of daily THI from January - March observations for a low year (red), an average year (green), and a high year (blue). The second panel shows the distribution of daily THI from January - March in a low year from observations (black line), unconditioned weather generator runs (small dashed line), and conditioned weather generator runs (large dashed line). The third panel shows the same as the second but for an average year, and the fourth panel shows the same as the second, but for a high year. In the low THI year (second panel) the peak of the PDF from the conditional simulation is closer to that of the PDF of the historical THI in this year and they both are shifted to the lower side. The unconditional simulations suggest higher values of THI with the peak of PDF shifted to the right. Similarly, in the high THI year of 2009 (bottom panel) the conditional simulation shows two peaks with the peak at higher value consistent with that of the historical. The unconditional PDF has a single peak at the lower end. These clearly demonstrate the effect of conditioning in effectively capturing the distribution of heat index.

Table 4-1. Observed and simulated THI exceedances for January - March. The table shows the number of days exceeding THI thresholds during this season for low, normal, and high years. Min. THI is the minimum number of exceedances from the 200 weather generator simulations and Max. THI is the maximum number of exceedances from the 200 weather generator simulations.

Year	Year Type	Dataset	THI >= 72	THI >=79
1989	Low	Observed	0.303	0.026
1989	Low	Unconditioned	0.433	0.067
1989	Low	Seasonal Conditioning	0.322	0.000
2004	Normal	Observed	0.484	0.077
2004	Normal	Unconditioned	0.440	0.065
2004	Normal	Seasonal Conditioning	0.418	0.000
2005	High	Observed	0.625	0.091
2005	High	Unconditioned	0.433	0.067
2005	High	Seasonal Conditioning	0.478	0.000

Table 4-1 shows the median probability of exceedance for two THI threshold levels for a low year, average year, and high year. This represents the number of days in a season exceeding the threshold. For THI > 72 exceedance probabilities more closely match the observed with conditioning. The probabilities are adjusted lower as compared to the unconditioned case in low years and adjusted high as compared to the unconditioned case in high years.

3.6.4 SIMULATIONS CONDITIONED ON SEASONAL CLIMATE FORECASTS

Conditioning covariates can be obtained from seasonal climate forecasts. These forecasts typically issue categorical probabilities – for example the probabilities of being above average, average, and below average. We demonstrate the utility of the conditional simulation approach in conjunction with seasonal climate forecast to help with planning decisions.

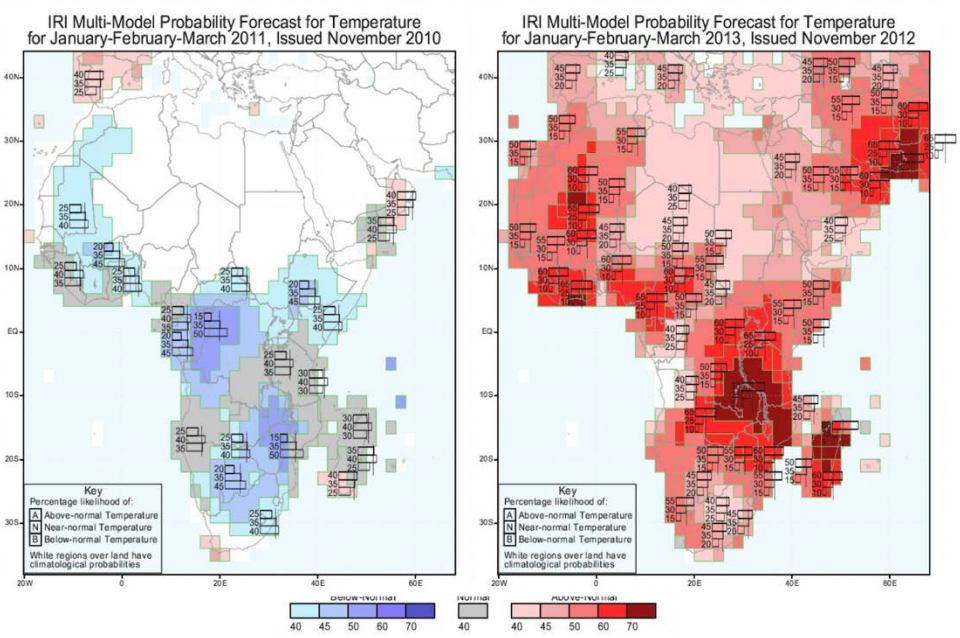

Figure 4-6. Examples of IRI seasonal forecasts for Africa. The left panel shows a forecast for a cool year and the right panel shows a forecast for a warm year. Bars show the probability of temperate being low, normal, or high.

The International Research Institute for Climate and Society (IRI; www.iri.columbia.edu) provides seasonal (three-month) probabilistic forecasts at one to four month lead-times. The forecasts are in the form of A:N:B likelihoods, where A is above-normal, N is near-normal, and B is below-normal. The categories are equal and defined with respect to climatological terciles (e.g., 33rd and 67th percentiles). The skill of these forecasts tends to be better at short lead-time. Figure 4-7 shows the seasonal forecast of temperature for the January - March season for 2011, a cool year, (left) and 2013, a warm year (right). First ensembles of seasonal historical temperatures are sampled with replacement based on the categorical forecasts. This is accomplished by categorizing the historical weather as above-, near-, or below-normal based on the empirical terciles, and then assigning the categorical forecasts as probabilities (or weights) to the historical values in each category and resampling with these assigned weights as the probability metric. We generated 100 ensembles of seasonal temperatures which were used as covariates to drive the weather generator 100 separate times. The output of these 100

independent runs is essentially a downscaled ensemble of weather patterns consistent with the seasonal climate forecasts. To investigate the sensitivity of THI to forecast skill we selected a below-normal (low) and an above-normal (high) year and assigned different probabilities of selecting a low or high year in the resampling. Results are shown in Figure 4-7 for a 60% certainty and an 80% certainty that an upcoming season will be either below-normal or above-normal.

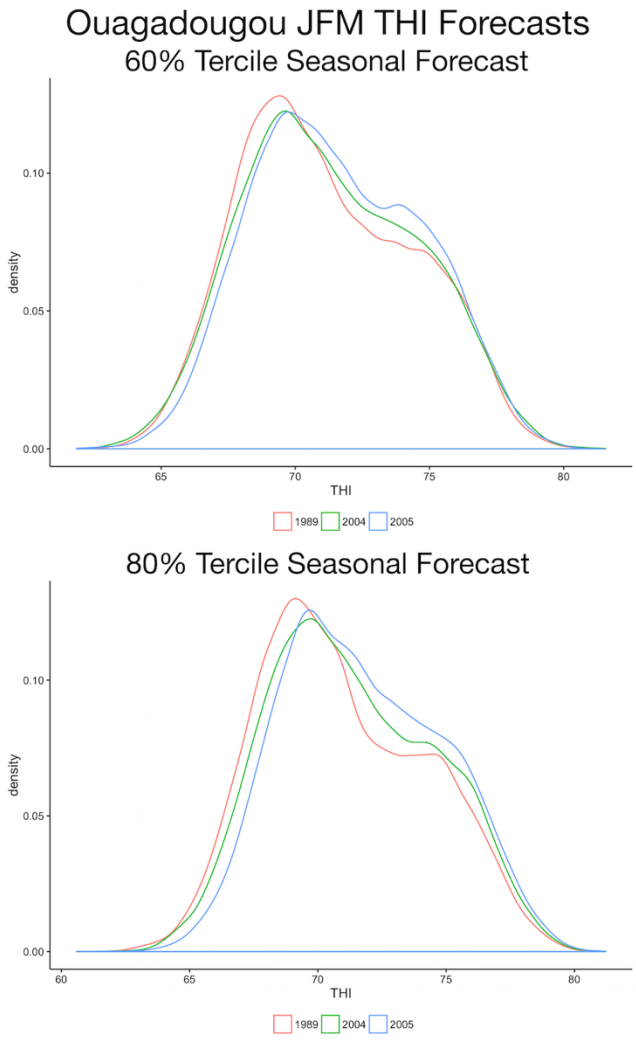

Figure 4-7. January - March THI low year (red), average year (green), and high year (blue) seasonally conditioned PDFs using a 60% certainty seasonal forecast (top) and 80% certainty seasonal forecast (bottom).

While the shifts in the THI PDFs are not as pronounced as seen in the observed data due to the introduced uncertainty it does move in the appropriate direction suggesting that using seasonal forecasts to condition the weather generator can improve the forecasting of THI.

3.7 DISCUSSION

Heat stress can have a large impact on livestock wellbeing, affecting their growth rate, milk production, and reproduction. It is especially relevant in regions where there is little control over the environmental exposure experienced by livestock. While heat stress is an animal-specific phenomenon, dependent on health, age, hide color, and access to fodder and water, a general understanding of the impact of different environmental conditions on heat stress can be found using heat indices and empirically derived threshold levels. Stochastic weather generators provide a method of producing complete weather sequences for evaluating heat stress indices. The incorporation of external covariates greatly improves the SWGs ability to capture the distributions of both the raw weather variables and the heat indices. They also offer the ability to provide seasonal forecasts or future projections when used with leading values or climate model output. In this paper, we only considered IRI seasonal forecasts and our synthetically generated forecasts which resampled the historical observations. Even greater improvements in skill would probably be realized when using more accurate seasonal temperature forecasts. These could be obtained from numerical weather prediction models, from statistical models that exploit persistence in large-scale climate, or a combination of the two. While it is possible to also use these forecasts to directly calculate heat stress measures such as the THI and HLI, incorporating them into a stochastic weather generator offers a way to obtain a daily time-series, as well as produce many realizations which can then be used to calculate probabilities of occurrence or probability of exceedance above a critical threshold. The space-time weather generator of Verdin et al. (2014) can be used to generate ensembles of heat stress

indices and consequently the health risks to livestock and enable effective mitigation strategies. The stochastic weather ensembles can also be coupled with physical models of livestock heat stress and simulate the indices. The demonstration of SWG for livestock heat stress application and also the extension to wind and relative humidity are unique contributions of this research.

4 SPATIAL AND TEMPORAL VARIABILITY OF EAST AFRICAN KIREMT PRECIPITATION AND LARGE-SCALE TELECONNECTIONS

4.1 INTRODUCTION

Ethiopia, north of about 7°N experiences one primary precipitation season, with a peak during the boreal summer, June to September. This season, termed the Kiremt, is responsible for 50%–80% of the annual precipitation over its agricultural regions and a majority annual precipitation falling in the headwaters of the Blue Nile and Omo Gibe rivers. Broadly, Kiremt season precipitation can be linked to two primary causes: the availability of moisture, and the behavior of local dynamical features promoting convection. There have been several studies investigating the sources and availability of moisture. One source identified is moisture transported from the Atlantic Ocean, Gulf of Guinea, and the Congo Basin by cross-continental low-level winds and delivered to the west side of the Ethiopian highlands (Flohn 1987). The strength of these low-level winds can be enhanced by a strong northeast directed pressure gradient between the Gulf of Guinea and the Arabian Peninsula. Segele et al. 2009a and Nicholson 2014 have found increases in Kiremnt precipitation associated with strenghtened low-level winds. Dry low-level easterly winds have also been shown to supress local convection which in turn depresses westerly moisture transport (Williams et al. 2012). Local circulation, which promotes convection can be largely linked to the East African Monsoon (EAM). The EAM is a low-level feature which establishes itself in the boreal spring and extends through the summer. The EAM is responsible for the development of the Tropical Easterly Jet (TEJ) and the low-level Somali Jet. Through increasing atmospheric instability, a strengthened TEJ is associated with increased precipitation, and a weakened TEJ is associated with decreased precipitation (Segele and Lamb 2005; Segele et al. 2009a,b; Diro et al. 2011; Williams et al. 2012). A strengthened low-level Somali jet is also associated with increases in precipitation (Nicholson 2014).

Sea surface temperature (SST) driven teleconnections have been linked to interannual variability of Kiremt season precipitation. At seasonal timescales,

teleconnections with El Nino Southern Oscillation (ENSO), the tropical Indian Ocean, and the tropical Atlantic Ocean influence moisture transport to this region the variability of the EAM, and thus Kiremt precipitation variability. The ENSO has been identified as a primary driver of interannual precipitation variability (Korecha and Barnston 2007). With El Nino conditions (warm central and eastern tropical Pacific) corresponding to periods of low precipitation and La Nina conditions (cooler central and eastern tropical Pacific) corresponding to periods of high precipitation (Nicholson and Kim 1997; Moran and Ward 1998; Seleshi and Zanke 2004; Diro et al. 2011). A weakening of the TEJ corresponds with El Nino conditions in the tropical Pacific (Chen and van Loon 1987; Nicholson 2013). Warmer conditions in the central and eastern tropical Pacific have been linked to an earlier monsoon onset, but a shorter season (Gissila et al. 2004; Block and Rajagopalan 2007; Korecha and Barnston 2007). In addition to ENSO, the Atlantic Ocean, and the Indian Ocean also influence Kiremt precipitation variability (Korecha and Barnston 2007; Nicholson 2014), with a warmer Indian Ocean and Arabian Sea seen to delay monsoon retreat and extend the precipitation season (Segele and Lamb 2005). At subseasonal scales, precipitation variability has been linked to the Madden-Julian Oscillation (MJO) via (1) the direct influence of the MJO core, (2) synoptic effects of the MJO core, and (3) influences on remote teleconnections via Rossby and Kelvin wave propagation (Zaitchik 2017). Direct MJO influences on East African long (March-May) and short (October-December) season precipitation have been noted. While there are subseasonal and spatial differences, these influences are the result of increasing westerly low-level moisture transport, weakened, but warmer and more moist low-level easterly winds, increases in low-level convergence, reduced stability, and increases in tropospheric air temperatures (Berhane and Zaitchik 2014). In addition to direct influences, there is an interplay between MJO and remote teleconnections including the Indian Ocean Dipole (IOD) and ENSO. There are findings that suggests MJO events preclude El Nino and have been linked to conditions important for its development. The MJO events during El Nino have been seen to have shorter lifecycles, though

they propagate further into the Pacific. While this influence has been seen to be strongest during the winter (Shimizu and Ambrizzi 2016), the influence during summer is not well understood. Enhanced MJO convection has been associated with the negative phase of the IOD with positive SST anomalies in the eastern Indian Ocean relative to the western Indian Ocean, and suppressed MJO convection has been associated with the positive phase of the IOD (Shinoda and Han 2005; Kug et al. 2009; Wilson et al. 2013; Benedict et al. 2015; Zaitchik 2017). Decreases in Kiremt precipitation occurred in concert with decreases in the West African Sahel during the 1970s and 1980s, however, while there has been a 'recovery' in the Sahel, precipitation in East Africa has continued to decrease in the recent 1990-present period (Williams et al. 2012). Furthermore, the teleconnections that drive Sahel rainfall show multidecadal variability (Chapter 3) and we suspect similar behavior with East African rainfall.

There have been several attempts to develop skillful seasonal precipitation forecasts for the Kiremt season using these teleconnections. Statistical regression approaches have demonstrated spatially-dependent skill at short lead times using large-scale ocean and atmospheric features. (Gissila et al. 2004; Korecha and Barnston 2007; Block and Rajagopalan 2007; Diro et al. 2011; Nicholson 2014). Predictability was strongly linked to the summer ENSO state, as well as discerning intensifying vs. decaying behavior in the spring. Skillful forecasts were limited to a two month lead time by the 'predictability barrier' of ENSO (Torrence and Webster 1998), rapid changes in circulation during the March-May season, and the sensitivity of the tropical Pacific and Indian Ocean to perturbations.

It is clear that to improve forecast skill of Kiremt season precipitation, understanding of the seasonal, subseasonal variability and potential predictability over time is crucial. Motivated by this need, in this paper, we perform analyses to quantify the temporal evolution of large-scale teleconnection on Kiremt precipitation. We also investigate connections with the MJO. Description of the data and methods are first presented followed by results and discussion.

4.2 DATA AND METHODS

Gridded precipitation data were obtained from the CRU TS3.23 dataset which provides monthly coverage for 114-year period of 1901-2014 at a 0.5°x0.5° resolution (Harris et al. 2014). Summer season (June-September) average precipitation along with subseasonal June (early season), July-August (peak season) and September (late season) average precipitation were computed. This dataset was selected for its relatively fine spatial resolution and length of record. We obtained daily precipitation from the Climate Hazards group Infrared Precipitation with Stations (CHIRPS) (Funk et al. 2015) dataset which offers a 0.05°×0.05° resolution from 1981-present. Global climate fields – winds, geopotential heights, sea level pressure, and precipitable water – were obtained from the NCEP/NCAR Reanalysis 1 (NNR) (Kalnay et al. 1996). The Nino 3,4 index was obtained from NOAA Earth Systems Research Laboratory (ESRL). Indian Ocean Dipole Mode Index (IOD-DMI), the Tropical North Atlantic Index (TNA), and the Tropical South Atlantic Index (TSA provides a timeseries of standardized SST anomalies over their respective regions and were calculated using the Hadley-OI merged SST dataset (Hurrell et al. 2008). The Nino 3,4 index bounded by 120°W-170°W and 5°S- 5°N (Trenberth 1997), IOD-DMI is defined as the temperature gradient between the East (90°E-110°E and 10°S-Equator) and West (50°E-70°E and 10°S-10°N) Indian Oceans (Saji et al. 1999). The TNA region is defined as 5.5°N-23.5°N and 57.5°W-15°W, and the TSA region is defined as 20°S-Equator and 30°W-10°E (Enfield et al. 1999). We constructed the Tropical Atlantic Index (TAI) as the difference between the TSA and TNA in order to capture the Atlantic dipole behavior.

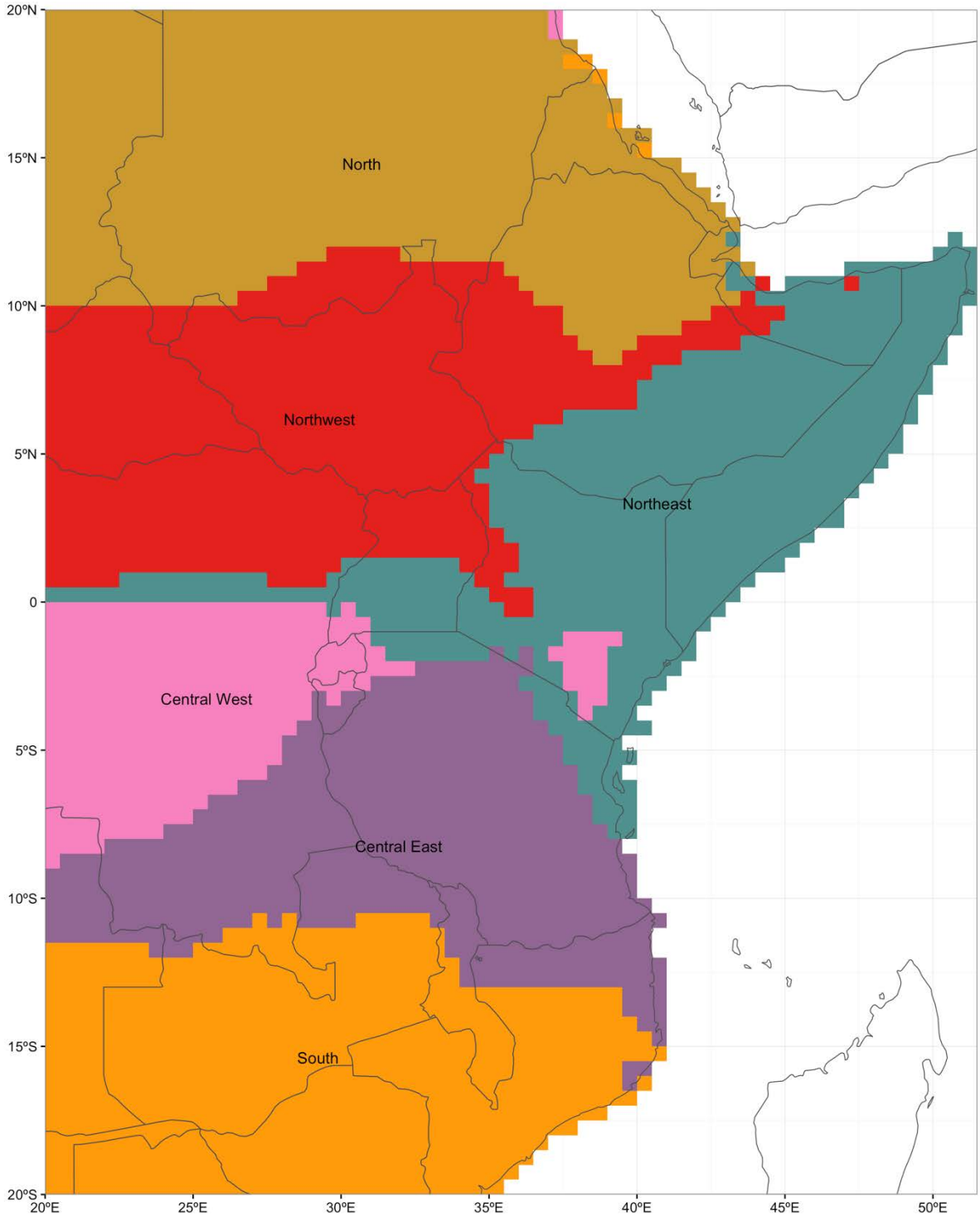

Figure 5-1. East Africa precipitation clusters, using a k-Means clustering approach on annual climatology and the CRU TS3.23 precipitation dataset

To identify regions of coherent rainfall regions, a K-means cluster analysis (Scott and Knott 1974) was performed on the annual climatology and the three-month periods of maximum precipitation at the stations. The climatology clusters the region into six distinct areas (Figure 5-1) shown in six different colors. The 'North' and 'Northwest' clusters have their primary rainfall season from June to September, i.e., Kiremt. The northern region in particular, contains the headwaters of several major rivers including, the Blue Nile and the Omo Gibe, and thus is the focus for this analysis. For the cluster, anomalous precipitation years are identified using a cluster-average standardized anomaly time series (1950-1999 climatology) with low (or dry) years corresponding to values <= -1 and high (or wet) years >= 1 for the full 114-year record. Composites of large scale climate fields for low and high years were produced for the full season as well as the subseasonal periods (early, peak, and late) as defined earlier. The fields selected are SSTs, low-level geopotential height (GPH) (850hPa), low (925hPa), mid (600hPa), and high (200hPa) level winds, and total-column precipitable water (precipitable water).

To examine the changes in strength of large-scale climate features, a Bayesian Dynamical Linear Model (BDLM) was employed. Traditional linear regression methods are incapable of capturing the changing relationship, i.e., nonstationarity, the BDLM, on the other hand, provides an attractive alternative to modeling and understanding the nonstationarity in the relationships (West and Harrison 1997; Petris 2009). BDLM has been applied to modeling Indian summer monsoon rainfall variability (Maity and Nagesh Kumar 2006; Krishnaswamy et al. 2015) and found interesting insights and better performance than the traditional regression methods showing the strengthening influence of IOD on Indian Monsoon and extreme rainfall events in recent decades while the influence of ENSO is weakening. Yanto et al. (2016) applied this to study the temporal variability of teleconnections of Indonesian rainfall with ENSO. In this method, the regression coefficients vary with time, unlike traditional regression where the coefficients remain fixed. A brief description is provided here for the benefit of the readers, for details see *West and Harrison*, 1997; *Petris,* 2009.

In BDLM the time series is considered as the output of a dynamical system perturbed by random disturbances (i.e., noise) – thus, considered a nonstationary evolution. The general form of the model is represented as:

$$y_t=F_t\ x_t+v_t\ v_t\sim N(0,V_t)$$
$$x_t=G_t\ x\_(t-1)+w_t\ w_t\sim N(0,W_t)$$

where y_t is a dependent variable (e.g. the leading PC), x_t is a vector of independent variables (e.g., Nino 3,4, TAI, etc.), t denotes time, F_t is an observation operator that transforms the model states into observations (e.g. regression coefficient) and G_t is the linear system operator (that models states of the independent variables). Both observation and system equations can have additive Gaussian errors with covariance matrices V_t and W_t.

The posterior predictive distribution of model coefficients θ_t at each time t is computed from the prior distribution of coefficients at time step t-1. Using Bayes theorem, the probability of the data yt conditional on the model parameters at time t is defined as:

$$P(\theta_t\ |y_t)\propto P(y\ y_t\ |\ \theta_t)P(\theta_t\ |y_{t-1})$$

The model coefficients θ consist of PCs, Nino 3,4, PDO, the regression coefficients and the variances of the Gaussian errors in the above equations. In the initial step, traditional linear regression is applied to compute regression coefficient with normally distributed mean and variance – i.e. $\theta_0\approx N(m_0,C_0)$ where m_0 and C_0 are the mean vector and the variance – covariance matrix of the regression parameters. Using a Kalman filtering approach together with Markov Chain Monte Carlo simulation approach, posterior distributions are estimated at any time t. The posterior distribution is then employed to generate the Bayesian confidence intervals (Petris 2009).

The BDLM returns transient regression coefficients for each covariate included, in this case the four SST indices, Nino 3,4, the TAI, and the DMI. These time-evolving coefficients can be compared to a fixed regression coefficient and its 5-95% confidence bounds (CI). Comparing the time evolution of these covariates and the changes in the standardized precipitation anomalies we were able to identify shifts in behavior. We used these shifts to define three epochs and repeat the above composite analysis.

4.3 RESULTS

Results focus on Kiremt season rainfall over the northern cluster (Figure 1) and are presented by analysis. Results are presented first for the seasonal and subseasonal composite analysis followed by the temporal variability of teleconnections analysis using BDLM, epochal composite results, and MJO composite results.

4.3.1 SEASONAL AND SUBSEASONAL COMPOSITE ANALYSIS

Composites of precipitation in high (i.e. wet) and low (i.e. dry) years are shown in Figure 5-2. Over the study region there are some notable subseasonal differences in precipitation anomalies. While the full season precipitation anomalies are uniformly wet and dry during the respective years (top row), early season drying (second row, right) is centered over east-central Ethiopia and northern Ethiopia, and there are slight positive rainfall anomalies seen in western Ethiopia. A similar pattern is seen during the peak season (third row, right). The late season drying pattern is more uniform, centered on central Ethiopia (bottom row, right). Wet patterns during the early season (second row, left) are more uniform, but the greatest increases shift north and east during the peak period (third panel, left). This wet/dry anomaly pattern can also be seen in the Sahel during the full season, with the clearest signal (homogenous spatial pattern) for both the wet and dry composites seen in the peak season (third panel). Similar wet/dry patterns can also

be seen in India (Gill et al. 2015a), with the most coherent pattern seen during the late season (bottom panel). These composite maps indicate that during wet and dry years there exists a coherent signal through tropical Africa and into India, indicative of global scale teleconnections.

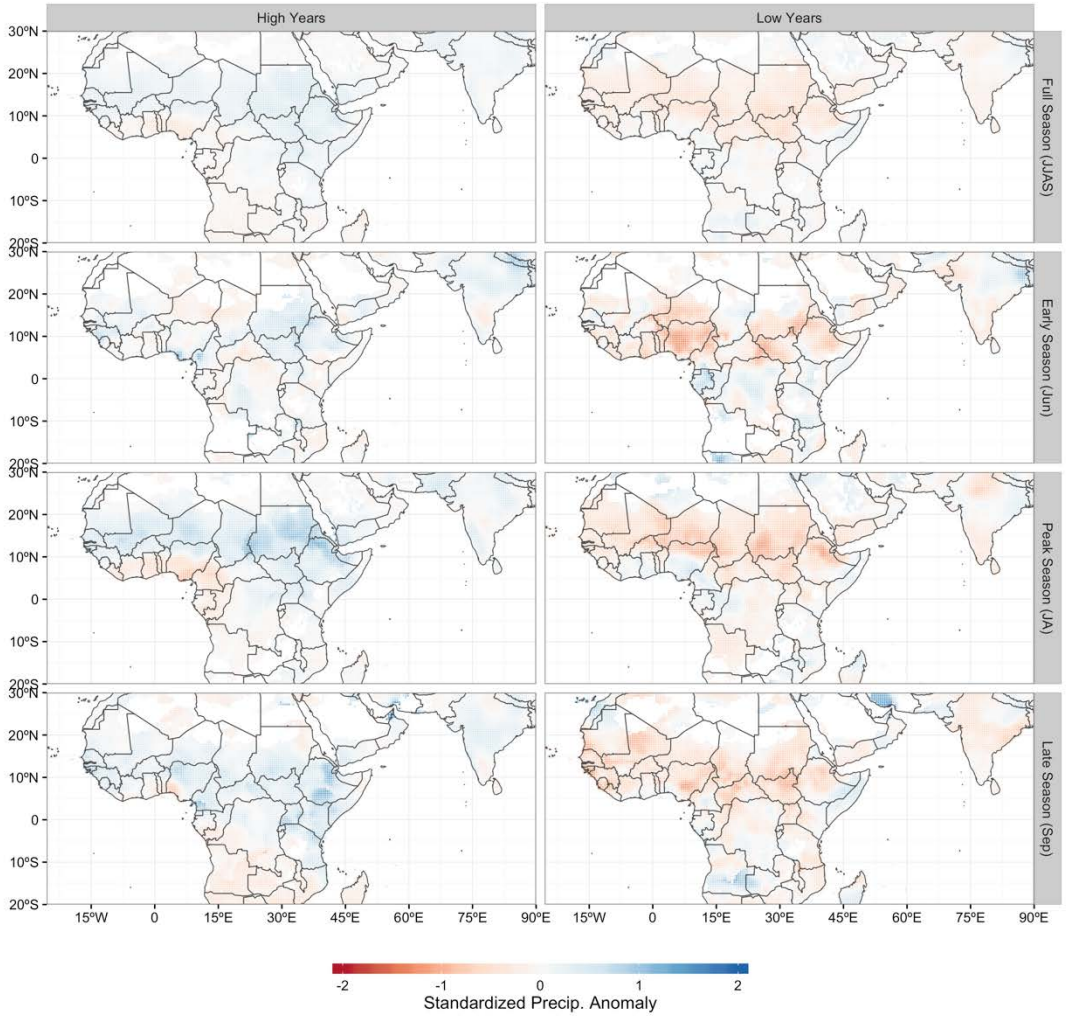

Figure 5-2. Composites of average precipitation corresponding to high (left column) and low (right column) precipitation years in the northern cluster based on the full record (1901-2014). Top row shows the composites for full season, June-September, 2nd row the early season, June, 3rd row peak season, July-August, and bottom row, late season, September.

In order to identify large-scale climate features existing during different precipitation regimes we first created composites of SSTs, low, mid, and upper level

winds, precipitable water, and low level GPH. These composites were created for the full season and for the early, peak, and late subseasonal periods for low and high precipitation years.

The SST composites for the high and low years and for the seasonal and subseasonal periods are shown in Figure 5-3. There are statistically significant differences between wet and dry years in the tropical Pacific, while these differences are not as pronounced in the tropical Atlantic. The tropical Pacific pattern, with cooler temperatures present in high precipitation years, and warmer temperatures present in low precipitation years is evident in all subseasonal periods. The Atlantic pattern, with cooler temperatures present in high precipitation years, and warmer temperatures present in low precipitation years is most apparent in the late season.

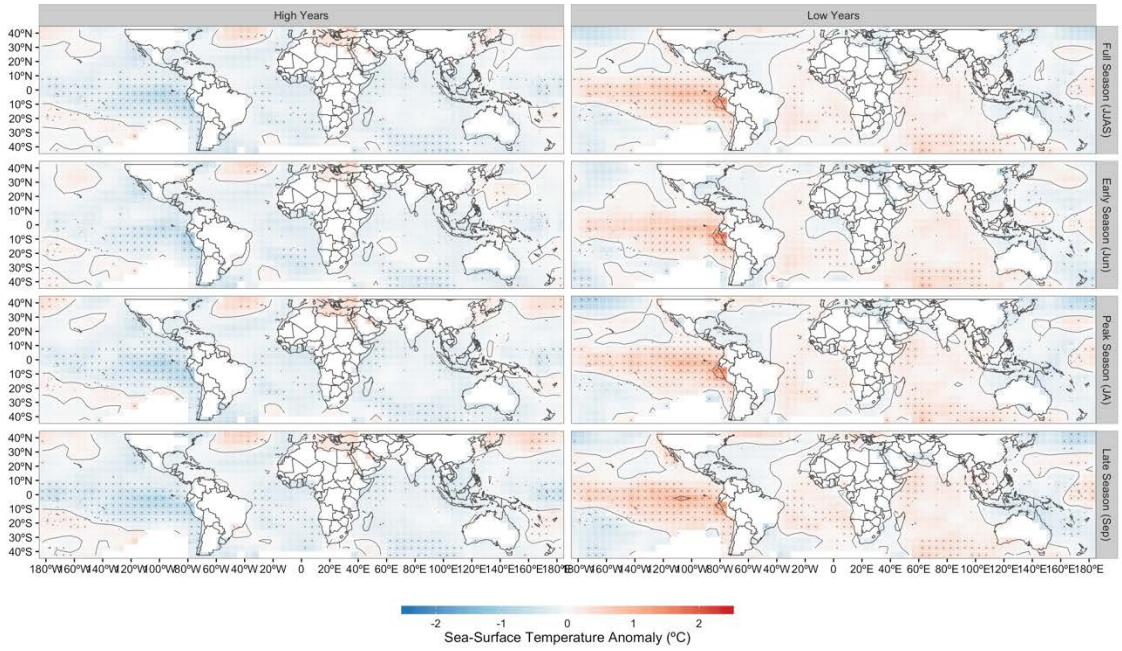

Figure 5-3. Same as Figure 2 but composites of SST anomalies. Statistically significant regions are stippled.

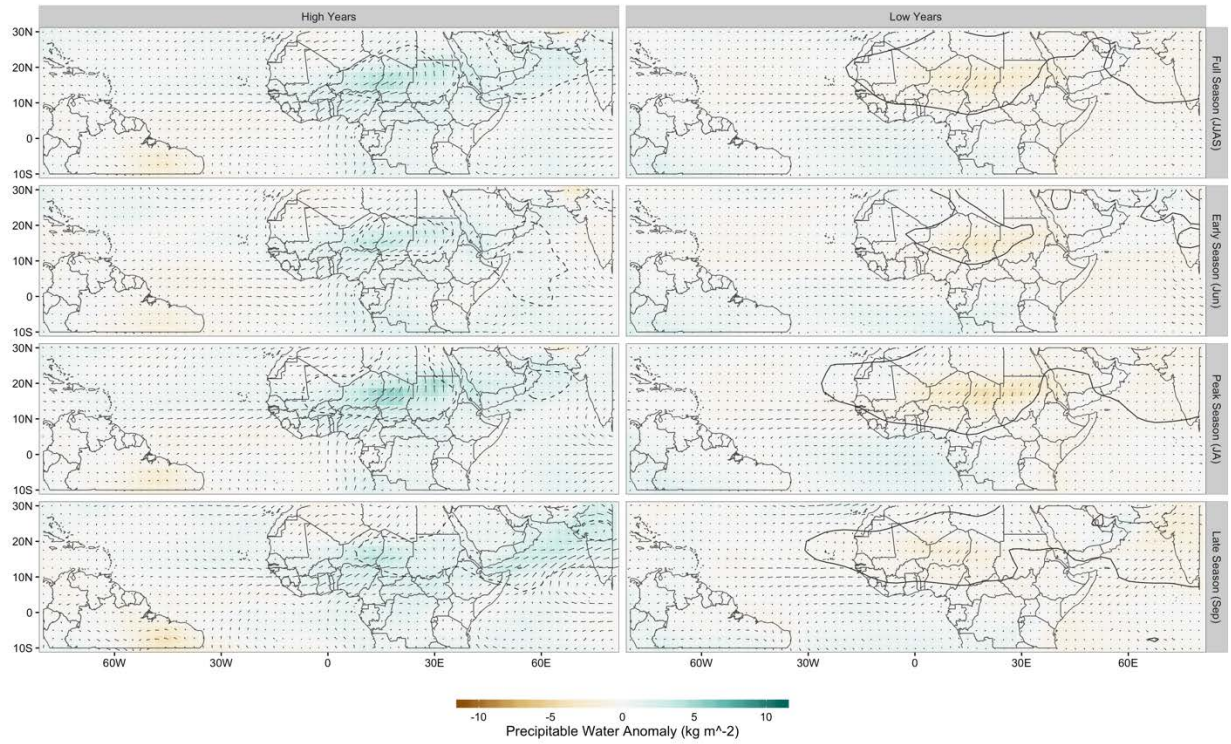

Figure 5-4. Same as Figure 2, but for average 925hPa winds (arrows), pressure (contours) and precipitable water (colors).

Composite anomalies of 925hPa winds, precipitable water, and 850hPa GPH are shown in Figure 5-4. The 925hPa winds correspond to the low-level monsoon winds that transport moisture into the region, and the 850hPa heights are selected to avoid surface friction. During high precipitation years, the cross-continental wind anomalies, originating in the Gulf of Guinea and passing over the Congo Basin are enhanced. These wind anomalies are most prevalent during the early and peak season, and relatively weaker during the late season. There are also wind anomalies extending from the southern Indian Ocean into Kenya and southern Ethiopia. These wind anomalies are consistent with the low-pressure anomaly over north central Africa, and the Arabian Peninsula. Conversely, during low precipitation years the cross-continental wind anomalies are absent, as are the southern Indian Ocean wind anomalies. This corresponds to a large region of high pressure anomalies extending across northern Africa and the Arabian Peninsula.

4.3.2 TEMPORAL VARIABILITY OF TELECONNECTIONS

The BLDM was applied to the summer season climate indices – TAI, IOD and Nino 3,4 and the summer seasonal average rainfall over the northern cluster. The time varying regression coefficients and the intercept are shown in Figure 5-5. All of the SST indices are known to influence precipitation during this season. Coefficients from the static (i.e. standard linear regression) regression and their 90% confidence intervals are shown as horizontal lines in these figures. Periods where the BLDM time-varying regression coefficients deviate outside the confidence bounds indicate it is significantly different and can be interpreted as changes in teleconnection strength.

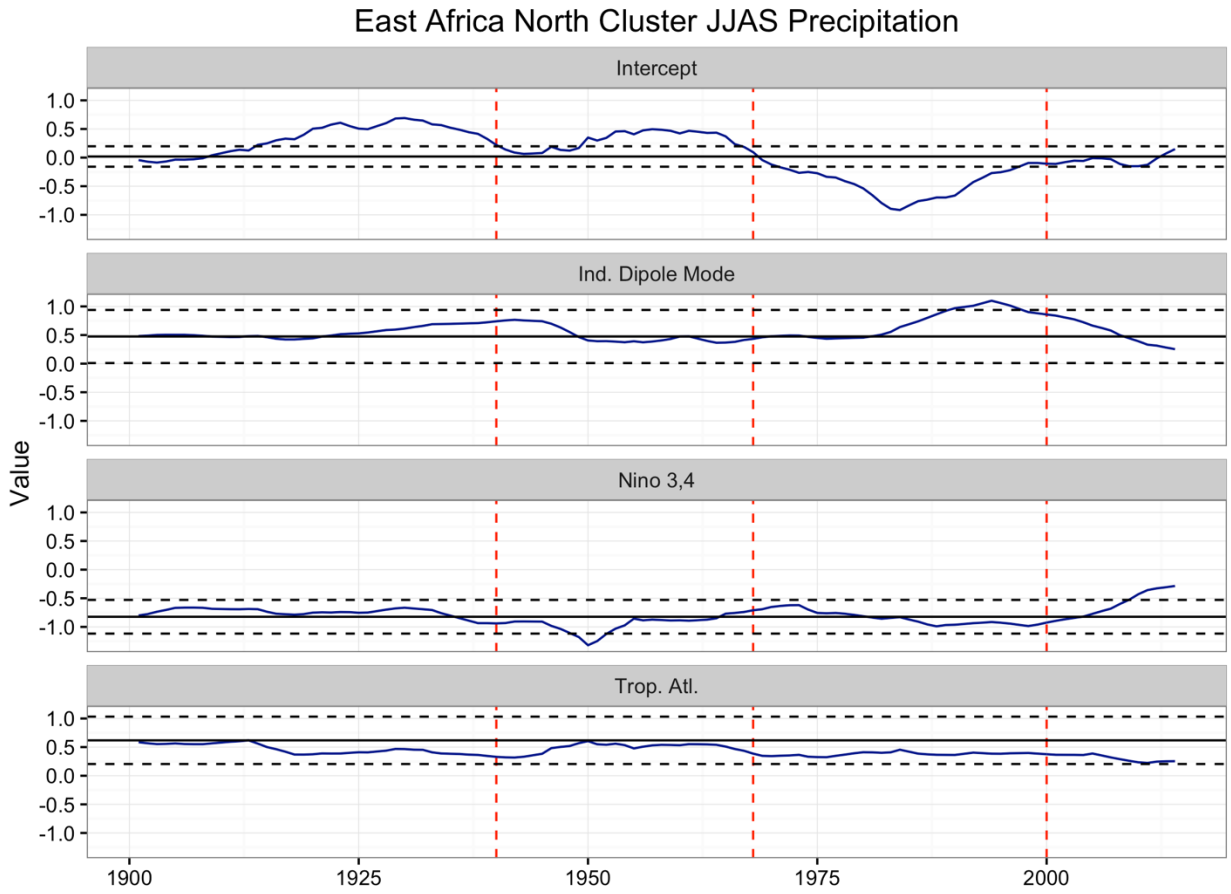

Figure 5-5. Time varying regression coefficients – intercept (top panel), IOD-DMI (2nd panel), Nino 3,4 (3rd panel), and TAI (bottom panel) – from BDLM for north cluster precipitation (blue lines). Coefficients from stationary linear regression are shown in black with 95% confidence intervals (dashed lines). Red lines show the boundaries of the defined epochs – wet period (1945-1968), drought (1969-2000), and modern period (2001-2014).

The intercept, which can be viewed as the mean value of the precipitation anomaly, shows positive values for most of the first half of the 20th century. From 1969 to 2000 precipitation anomalies were negative, reaching their lowest point in the 1980s, before returning to 0 in the modern period. This indicates weaker rainfall during the latter part of 20th century with modest recovery in recent years – consistent with the feature seen with Sahel rainfall (Broman et al., in review). The relationship with IOD-DMI shows no temporal variability and the values are within the interval from the static regression. In the recent decade, this relationship has

weakened (headed toward 0). The coefficient of Nino 3,4 has remained negative over the entire 114 years (negative relationship with ENSO), and also remained relatively constant. In recent decade, this relationship has also weakened, significantly in the last 10 years. The relationship with TAI is similar, in that a positive relationship has existed consistently over the whole record, though has begun to weaken in recent years. This analysis suggests that the teleconnections have remained fairly stable however, in recent years they are weakening. This is of concern in that it also suggests that the predictability might also be weakening.

Timeseries of the indices and standardized precipitation are shown in Figure 5-6. The top panel shows an anomalous wet period in East Africa through the late 1960s, followed by the dry epoch until the end of the century. The Nino 3,4 shows the well-reviewed multidecadal pattern of cooler during the middle part of the 20th century and warmer subsequently with a cooling in the early 21st century. The TAI also shows a multi-decadal pattern, and interestingly has been increasing from the 1990s to the present. The IOD-DMI is fairly constant through the entire period. The epochal shifts in precipitation correspond to shifts in the strengths of the SST teleconnections as seen in the BDLM regression coefficients (Figure 5-5**)** and also the temporal variability of the indices in Figure 5-6.

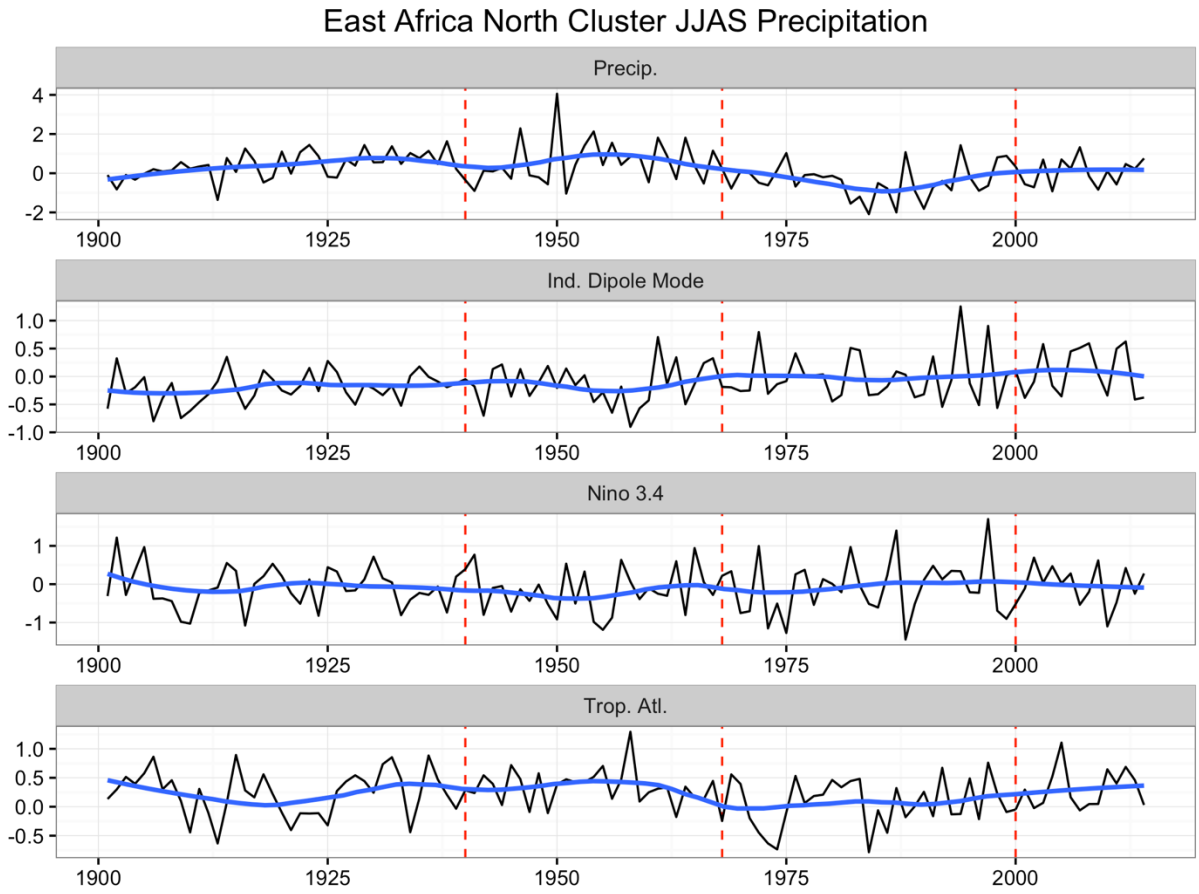

Figure 5-6. Standardized precipitation and SST index (black line) and smoothed values (blue line) for north cluster precipitation (top panel), IOD-DMI (2nd panel), Nino 3,4 (3rd panel), and TAI (4th panel). Red lines show the boundaries of the defined epochs: wet period (1945-1968), drought (1969-2000), and modern period (2001-2014).

4.3.3 EPOCHAL COMPOSITES

For the epochal composites, we have lowered the threshold for identifying low and high precipitation years to +/- 0.4 in order to identify low years during the mid-century wet period and high years during the late-century drought.

During the mid-century wet period (Figure 5-7), there are significant differences between

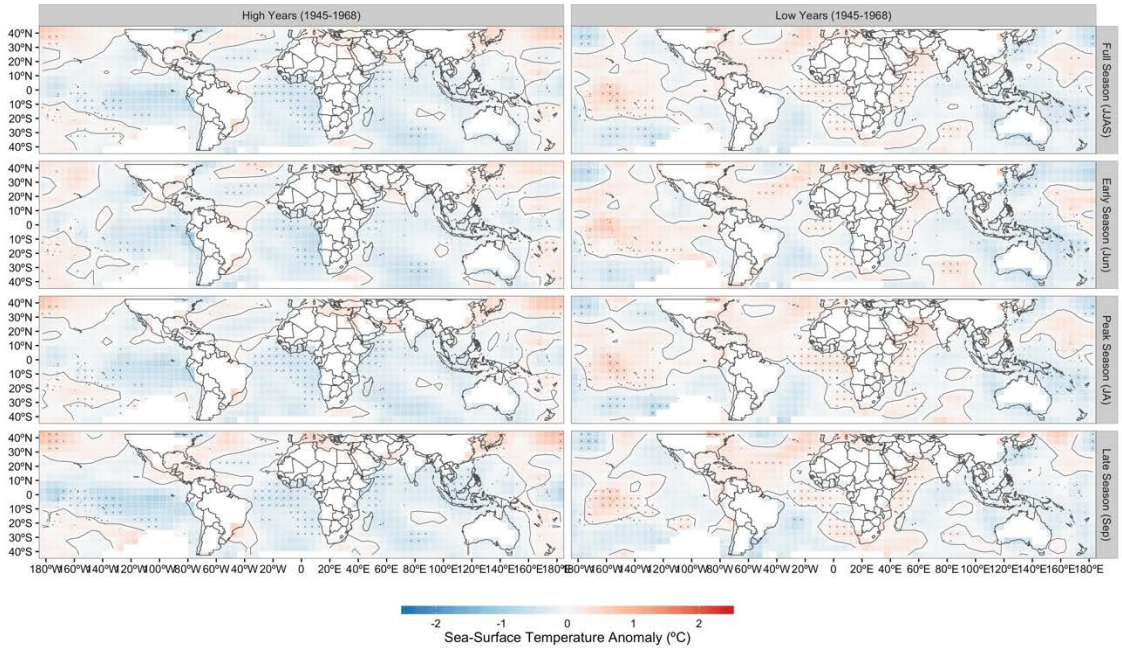

Figure 5-7. Same as Figure 3, but for the wet period (1945-1968)

the high and low precipitation years in the Atlantic and Pacific. These anomalies are clearly seen in the full season (Figure 5-7, top row) where in, during high years the entire tropical oceans (Atlantic, Pacific and Indian) are cooler and during low years, they are warmer but less coherent. Subseasonally they appear at the peak and late seasons (Figure 5-7, 3rd row and bottom row). There is a reduced influence of the Pacific, where negative anomalies (La Nina-like features) are seen during high precipitation years, but the corresponding positive anomalies are weaker during low precipitation years.

The decline in precipitation in East Africa in the late-20th century corresponded to similar declines in precipitation over the West African Sahel. SST composites for low and high precipitation years during this epoch (Figure 5-8) show a clear ENSO signal, but warmer in Atlantic and Indian Oceans during both high and low years. In the tropical Pacific, La Nina conditions are present during wet years and El Nino conditions are present during dry years. This ENSO signal is seen for the full season and for all three subseasonal periods. There are no statistically significant differences between SST anomalies in the Atlantic or Indian

Oceans. Interestingly, these ocean regions have negative anomalies for the full 1901-2014 period (Figure 3).

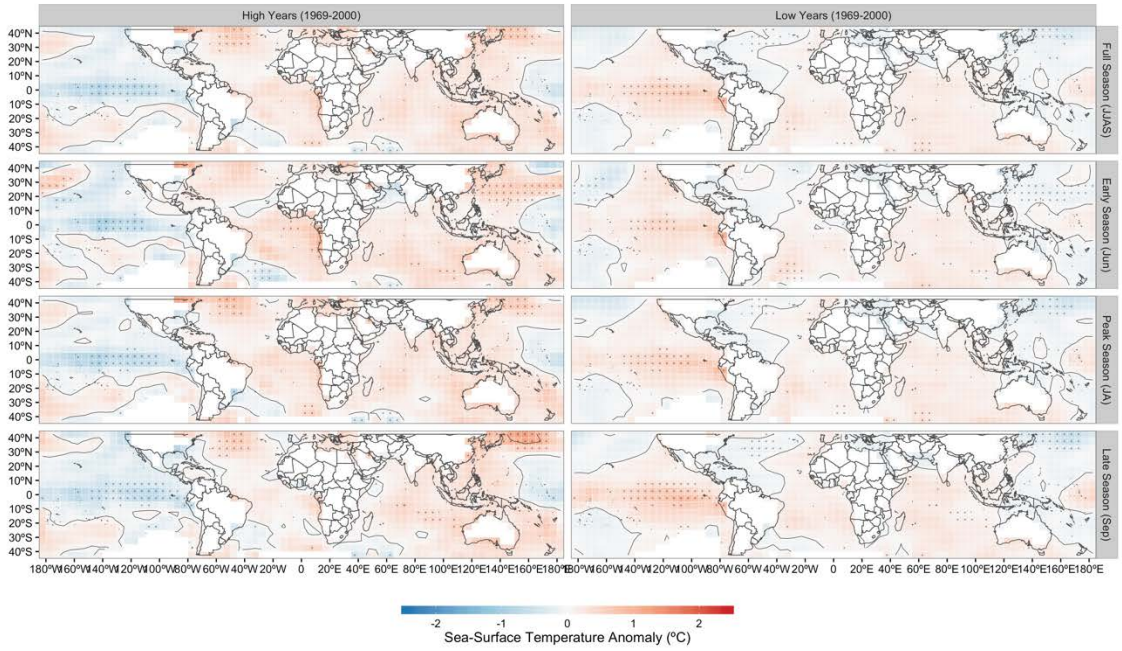

Figure 5-8. Same as Figure 8, but for the drought period (1969-2000)

SST anomalies in the modern period (Figure 5-9) show an increase throughout for both wet and dry years. A consequence of this warming has been a decrease in the strength of remote teleconnections, including ENSO, the Indian Ocean, and the Atlantic Ocean.

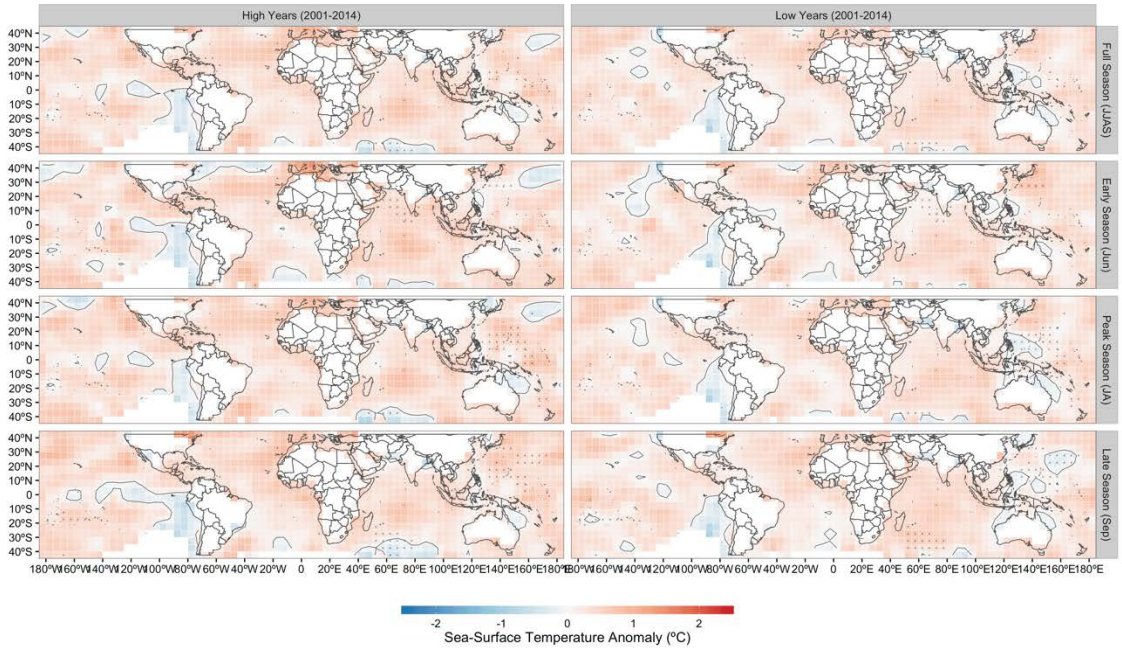

Figure 5-9. Same as Figure 8 but for the modern period (2000-2014).

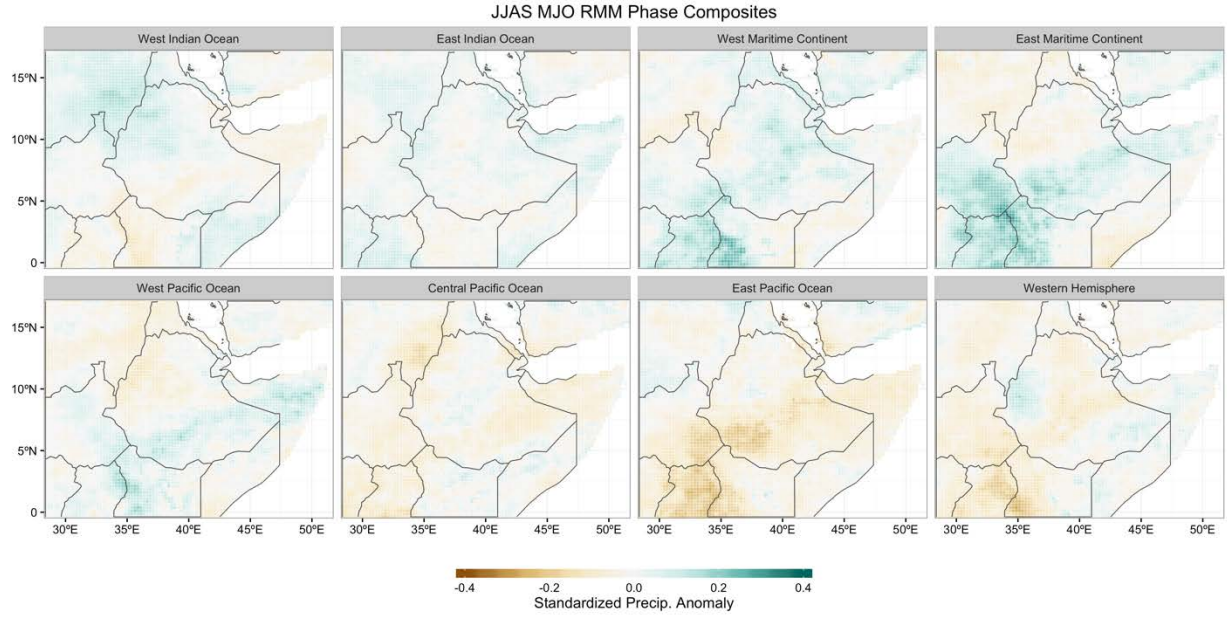

Figure 5-10. Composites of June-September average precipitation anomalies (using CHIRPS) corresponding each phase of the MJO. Periods with no MJO activity are not shown.

4.3.4 MADDEN-JULIAN OSCILLATION COMPOSITES

To examine the influence of the MJO on June-September precipitation, a similar composite analysis was performed using MJO phase. Figure 10 shows composites for June-September precipitation during each phase of MJO. Weak increases in precipitation are seen during the 'West Maritime Continent' phase and more strongly during the 'East Maritime Continent' phase. The strongest influence appears during the 'East Pacific Ocean' phase where there are negative precipitation anomalies across a majority of Ethiopia. While these anomalies are not as strong or coherent as seen in the spring or fall precipitation seasons (Berhane and Zaitchik 2014; Zaitchik 2017), they do indicate an MJO influence on precipitation during the summer season.

4.4 DISCUSSION

In this chapter, we investigated the interdecadal and intraseasonal behavior of East African summer monsoon precipitation and the corresponding variability in SST teleconnections and large-scale dynamical fields. Over the past century the wet and dry precipitation anomaly pattern seen in East Africa corresponds with anomalous conditions in the Sahel and in India. This suggests that variability in the large-scale teleconnections influencing precipitation can influence tropical precipitation on a large scale. The subseasonal signature of this correspondence is strongest in the peak season for the Sahel and in the late season for India. Similar relationships can be seen when looking at composites of precipitation during La Nina and El Nino years. As with the Indian monsoon, there is an anomalous influence with the increases in precipitation seen in La Nina years showing a much stronger and spatially homogenous signal as compared to the decreases in precipitation seen in El Nino years.

The strength in the relationship between these large-scale teleconnections and precipitation has varied over the past century as shown from the BDLM results. Over the last century the relative strengths of the tropical Pacific, the

Indian Ocean, and the tropical Atlantic have varied slightly, with ENSO strengthening during the mid-century wet period, and the Indian Ocean strengthening during the second half of the late-century drought. In the modern period, increases in global SSTs have further altered this relationship, with the Indian Ocean, ENSO, and the tropical Atlantic all weakening. Circulation feature anomalies clearly show the large-scale dynamics in the region including the low-level cross-continental moisture transport.

At shorter timescales we have shown that MJO activity can influence precipitation, with the most coherent signal being drying during the 'East Pacific Ocean' phase. Past studies have shown strong connections between the MJO and precipitation in the spring and fall seasons. Here we show that, while weaker, the MJO also influences precipitation during the summer season.

Collectively these analyses indicate that large-scale teleconnections influence precipitation during the summer season, however, there has been a weakening of the relationship in the early 21st century. Results focused on northern Ethiopia, the headwaters region for several major rivers including the Omo Gibe and Blue Nile. This region receives a majority of its precipitation during the Kiremt season, and being this headwaters region, an understanding of precipitation variability and development of skillful forecasts is important in water management. These teleconnections have been used in past studies to develop seasonal precipitation forecasts, however with the understanding that there has been a shift in these relationships in recent decades, care should be taken when developing future forecasts. Identifying the MJO as influencing precipitation during this season now offers the possibility of incorporating it into new forecasting techniques. Chapter 6 discusses the development of new statistically-based forecasts, and post-processing of dynamical forecasts using the findings presented in this chapter.

5 S2S PRECIPITAITON FORECASTING FOR WATER MANAGEMENT IN ETHIOPIA

5.1.1 INTRODUCTION

The highlands of Ethiopia, north of about 7ºN, experience one primary precipitation season termed the "Kiremt" (June through August). In this region, 50-80% of the annual total precipitation over agricultural regions fall during this season. A majority of the headwaters precipitation for the Omo Gibe and Blue Nile rivers also fall during this season. Skillful forecasts of precipitation during the subseasonal to seasonal timescale (S2S), typically, two weeks to two months, would fill the current gap between currently available short-term weather forecasts and longer-term seasonal forecasts. Skillful S2S forecasts offer the water management community the ability to anticipate longer-term water supply demands and adjust water management operations accordingly, including setting reservoir levels, setting reservoir releases, and optimizing hydropower production. The uncertainty of future conditions can also be quantified when properly-calibrated ensembles of these forecasts are produced, allowing water managers to make operational decisions based on a desired risk tolerance (Sankarasubramanian et al. 2009; Webster 2013; Friend et al. 2014).

On the seasonal timescale, there have been several attempts to develop skillful precipitation forecasts for the Kiremt season which can offer insights when developing a new forecasting approach at the S2S timescale. Statistical regression approaches have demonstrated spatially-dependent skill at short lead times using large-scale ocean and atmospheric features. (Gissila et al. 2004; Korecha and Barnston 2007; Block and Rajagopalan 2007; Diro et al. 2011; Nicholson 2014). Predictability was strongly linked to the summer ENSO state, as well as discerning its intensifying versus decaying behavior in the spring. Skill in forecasts were limited to a two month lead time by the 'predictability barrier' of ENSO (Torrence and Webster 1998), rapid changes in global circulation patterns during the March-May season, and the sensitivity of the tropical Pacific and Indian Ocean to

perturbations. Shukla et al. 2016 investigated the seasonal predictive skill in numerical weather prediction (NWP) forecasts using the North American Multimodal Ensemble (NMME) (Kirtman et al. 2014) and found limited deterministic skill beyond a two-month lead time. Rank probability skill score (RPSS) was higher in active ENSO years (El Nino or La Nina), which corresponds well with the strong Kiremt precipitation-ENSO teleconnection discussed in Chapter 4 and in the statistical forecasting attempts discussed above.

The S2S timescale has been given increased interest in recent years for its ability to provide information at timescales critical for decision making (Robertson et al. 2015). A suite of post-processing approaches that have been applied to other forecasting applications (Sankarasubramanian and Lall 2003; Hopson and Webster 2010a; Towler et al. 2010; Webster et al. 2010) were applied to raw S2S forecasts for basin-wide precipitation in Ethiopia. As discussed in Chapter 4, the Madden-Julian Oscillation (MJO) has been shown to have influence on seasonal precipitation. In the analysis we preset in this chapter we use the MJO in post-processing to condition the forecasts. The influence of the post-processing approaches along with the influence of incorporating in the MJO on S2S forecast skill will be evaluated and presented below.

5.2 DATA AND METHODS

Precipitation forecast data were obtained from the National Center for Environmental Prediction (NCEP) Climate Forecast System (CFSv2) (Saha et al. 2014) which has a grid resolution of roughly 100km (T126) and is available for the period 2011-2016. Four ensemble members are available each day extending out 120 days; the 0-60 day forecasts were retained for this analysis. Reference precipitation data against which we evaluate the forecasts were obtained from the Climate Hazards group Infrared Precipitation with Stations (CHIRPS) (Funk et al. 2015) dataset which blends satellite precipitation estimates with observation stations at a 0.05°×0.05° grid resolution and for the period 1981-present. The MJO can be quantified by its anomaly and phase (Wheeler and Hendon 2004); MJO index

data were downloaded through the IRI data library (http://iridl.ldeo.columbia.edu/) and originated from the Australian Bureau of Meteorology dataset.

Forecasted precipitation and reference precipitation were basin-averaged to calculate timeseries of daily precipitation over each of the twelve river basins shown in Figure 6-1.

Figure 6-1 River basins used for precipitation forecasting. Mean precipitation rates were calculated for each basin from both the CFSv2 forecast data and the CHIRPS gridded data.

These basin-average precipitation timeseries were then aggregated to different accumulation periods – 24hr, 5day, 10day, 15day, 20day accumulations. This was done to examine the dependence of forecast skill on temporal integration length.

Finally, to focus only on the precipitation season, months with less than 5% of the average annual precipitation for each river basin were removed. The basin-average forecasted precipitation timeseries are termed the 'raw dynamical' forecasts. A second set of dynamical forecasts was obtained by applying a quantile mapping (Q-Q) bias-correction approach and quantile regression (QR) post-processing. A final set of dynamical forecasts was obtained by applying the same bias-correction and post-processing approaches, but using the MJO anomaly and phase to condition the forecasts. Thus, we have three sets of dynamical forecasts for each river basin and for each accumulation period:

a) raw forecasts

b) bias-corrected and post-processed forecasts

c) bias-corrected and post-processed forecasts conditioned on the MJO anomaly and phase

The bias-correction and post-processing approaches are described in detail below.

5.2.1 BIAS-CORRECTION AND POST-PROCESSING

The raw dynamical CFSv2 forecasts were first bias-corrected to the CHRIPS basin-averaged precipitation using a Q-Q approach which adjusts the entire empirical PDF of the forecasts to match the empirical PDF of the CHIRPS (Hopson and Webster 2010b). Following Q-Q bias-correction, QR post-processing was then applied. QR has been used to post-process precipitation forecasts including in studies by Bjørnar Bremnes 2004, Hopson and Webster 2010, and Towler et al. 2010. QR is a type of regression that seeks to conditionally fit quantiles of a regressand's distribution while making no assumption of the underlying distribution of the regressand or error.

After bias-correction, QR was applied to the forecasts. In applying QR to ensemble forecasts, the number of quantiles is determined by the desired number of ensemble members, with the minimum and maximum quantiles assigned to the lowest and highest ensemble members. For example, with an evenly-spaced three-

member ensemble, the lowest member would represent the 25th percentile, and the highest member would represent the 75th percentile, and the middle member would represent the 50th percentile, however the actual percentiles chosen can be arbitrary The QR approach is as follows. Assume $\{y_i\}$ is a set of observations of the regressand y, and $\{\mathbf{x}_i\}$ is an associated set of predictors. Just as in standard linear regression, a linear function of $\mathbf{x}$ can be used to estimate a specific quantile q_θ of y:

$$q_\theta(\mathbf{x}_i;\boldsymbol{\beta}) = \beta_0 + \sum_{k=1}^{n} \beta_k x_{ik} + r_i$$

with residual $r_i = y_i - q_\theta(\mathbf{x}_i;\boldsymbol{\beta})$ and $\theta \in (0,1)$, wherein $\boldsymbol{\beta}$ is a vector of unknown coefficients. However, instead of minimizing the squared residuals, as done with standard linear regression, in QR a weighted iterative minimization of $\{r_i\}$ is performed over $\boldsymbol{\beta}$:

$$\min \sum_{i=1}^{n} \rho_\theta(r_i) = \arg\min_{\beta} \sum_{i=1}^{n} \rho_\theta(y_i - q_\theta(\mathbf{x}_i;\boldsymbol{\beta}))$$

with weighting function

$$\rho_\theta(u) = \begin{cases} \theta u & u \geq 0 \\ (\theta - 1)u & u < 0 \end{cases}$$

In applying QR to the CFSv2 precipitation forecasts, the ensemble mean, ensemble quantile timeseries, and persistence forecasts were used as predictors in the initial set. The ensemble median was calculated as the 50th percentile forecast from the multiple forecast ensemble members. Ensemble quantile timeseries forecasts were calculated at evenly spaced quantiles as described above by interpolating between the ranked forecast ensemble members. Finally, persistence forecast were assembled by using the current reference precipitation value at the initialization date as the forecast of precipitation at the desired lead-time. These approaches were used to develop the second set of dynamical forecasts after applying cross-validation in the selection of the final regressor set.

In developing the third set of dynamical forecasts, we added to the set of regressors in QR timeseries developed using a forecast developed from using the MJO. The MJO anomaly and phase can be used to develop statistically-based naïve precipitation forecasts (described below). From these forecasts timeseries of precipitation anomaly and precipitation standard deviation can be obtained and these are added to the CFSv2 ensemble mean, CFSv2 ensemble quantile timeseries, and persistence forecast in QR. Development of this naïve MJO forecast is described below.

5.2.2 NAÏVE MJO FORECAST

As discussed in Chapter 4, the MJO has been shown to influence precipitation in East Africa on the months to seasonal timescale. This suggests that including a measure of the MJO in the post-processing of dynamical precipitation forecasts could improve forecasts at the S2S timescale by shifting the ensemble higher during MJO phases that correspond with wetter conditions and shifting the ensemble down during MJO phases that correspond with drier conditions. The MJO is not always active, but follows a predictable pattern when it is active. To take advantage of this predictability we developed a 'naïve MJO forecast' as the MJO measure used to post-process dynamical precipitation forecasts. We term it a 'naïve' forecast as it relies solely on the historical behavior of precipitation during a particular MJO phase and is not informed by any other measure of local conditions. The thought behind this naïve forecast is MJO influence on precipitation when the MJO is active is fairly consistent, and while it may not be able to provide detailed S2S precipitation forecasts, it can provide reasonable information on the sign and magnitude of change. When used as a regressor in QR this information can adjust the forecast ensembles in the correct direction and improve forecast skill. Exploiting the ideas that the MJO follows a predictable pattern when active and that its influence on precipitation under different MJO phases is consistent, we developed the naïve MJO forecast as the mean precipitation total over a given accumulation

period using historical observations, given the current MJO phase. This is to say if we wish to issue a forecast today for week 4 of the month and the MJO is in phase 3, using observation, we find all of weeks where three weeks previously the MJO was also in phase 3, and take the mean of accumulated precipitation from all of the identified weeks. One further step is take to refine this mean accumulated precipitation and arrive at the naïve MJO forecast – we remove the climatological mean accumulated precipitation which leaves precipitation anomalies, the signal of MJO influence on precipitation. This mean accumulated precipitation anomaly is our naïve MJO forecast and assumes that whenever the MJO is in phase 3, in three weeks' time, because it is predictable, will be in a consistent phase, and that when the MJO is in that phase, it influences precipitation in a consistent manner. In order to account for seasonal differences in how the MJO influences precipitation, when identifying the set of weeks to average, only weeks within the same season are considered. For each basin, seasons were defined using the daily climatology – season definitions try and capture the main precipitation period. In some basins, there were multiple precipitation seasons; for these a season is defined for each period. Season definitions for the Abbay (left) and the Omo Gibe (right) are shown in Figure 6-2.

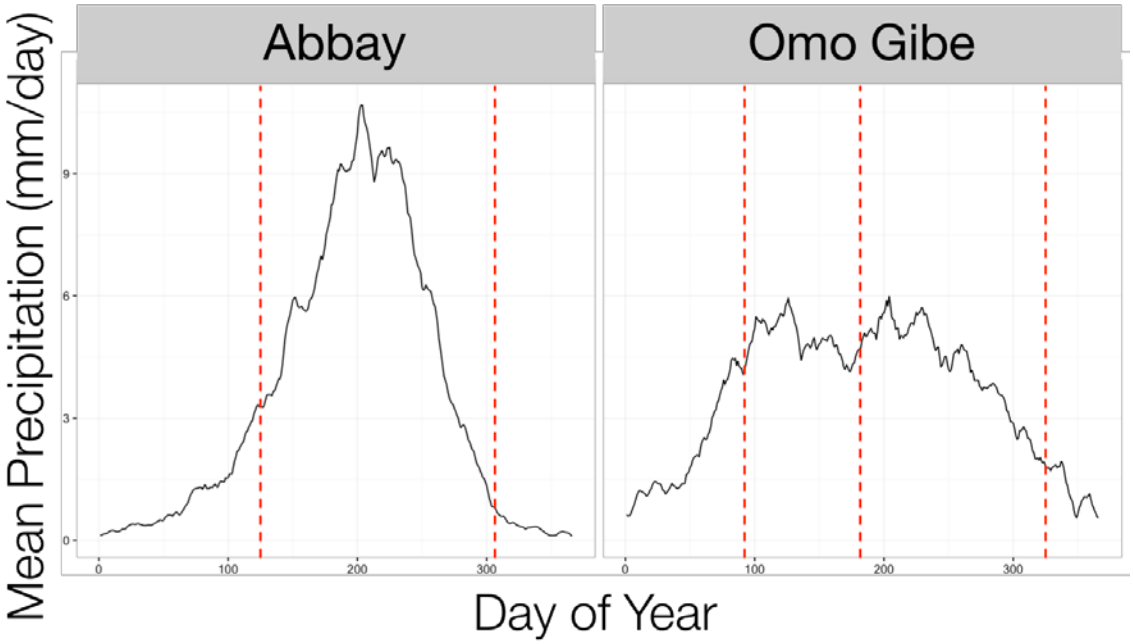

Figure 6-2 Average Precipitation Daily Climatology using CHIRPS for the Abbay river basin (left) and the Omo Gibe river basin (right). Red lines show the season bounds.

The Abbay basin being further north experiences one primary precipitation season, with peak precipitation occurring in the summer. For this basin, two seasons were defined, a dry season from November to April, and a wet season from May to October. The Omo Gibe basin, by contrast, straddles the summer-dominant precipitation region, and the spring-fall dominant precipitation region. To account for potential differences in precipitation dynamics, two primary precipitation seasons were defined: one from March to June, and a second from July to November, along with a dry season from December to February. Naïve MJO forecasts were constructed separately for each precipitation season; one forecast was constructed for the Abbay basin for the May to October season, and two forecasts were constructed for the Omo Gibe basin: one for the March to June season and another for the July to November season. One final piece of information can be teased out from this approach – as multiple historical periods go into calculating the mean accumulated precipitation, we can use the same historical observations to calculate the standard deviation of accumulated precipitation. This

quantifies the uncertainty in the influence of the MJO on precipitation and can also be used in QR post-processing. Precipitation anomaly and precipitation standard deviation from the naïve MJO forecasts were incorporated as covariates and used to develop the third set of dynamical forecasts within the QR post-processing step, as explained above.

An initial assessment of the CFSv2 forecasts was performed by comparing the daily mean climatology of the forecasts against the daily mean climatology of the CHIRPS basin-average precipitation. This shows how well the NWP forecast model represents the large-scale dynamics and seasonal climatology of this region. Forecast skill was then evaluated as follows: Raw CFSv2 forecasts were correlated with the CHIRPS basin-average precipitation and monthly-average correlations for different forecast lead times were calculated. Each set of dynamical forecasts, raw, bias-corrected and post-processed, and bias-corrected and post-processed conditioned on the MJO were then compared against climatology and persistence in their ability to forecast precipitation. The CHIRPS basin-average precipitation were used as the reference or observed precipitation. Briar Skill Scores (BSS) were calculated for each set of dynamical forecasts and for each accumulation period, 24hr, 5days, 10days, 15day, and 20days, with forecasts evaluated over the period 0 to 60 days. Results are presented for the Omo Gibe and Abbay River basins and provide an assessment of the skill of the CFSv2 at the S2S timescale as well as the improvements gained from bias-correction and post-processing. Forecast improvements incorporating in the MJO into the QR post-processing are also shown in the figure, discussed below.

5.3 RESULTS

To first examine the skill of the NWP forecast model in representing the large-scale dynamics of Ethiopia, daily climatologies of the CFSv2 forecasts and CHIRPS reference precipitation were compared. Figure 6-3 shows the observed and forecasted daily climatologies for 5-day accumulations. The CFSv2 forecasts match

the seasonal climatology of reference precipitation well, capturing the single precipitation peak seen in the Abbay basin and the double peak seen in the Omo Gibe basin.

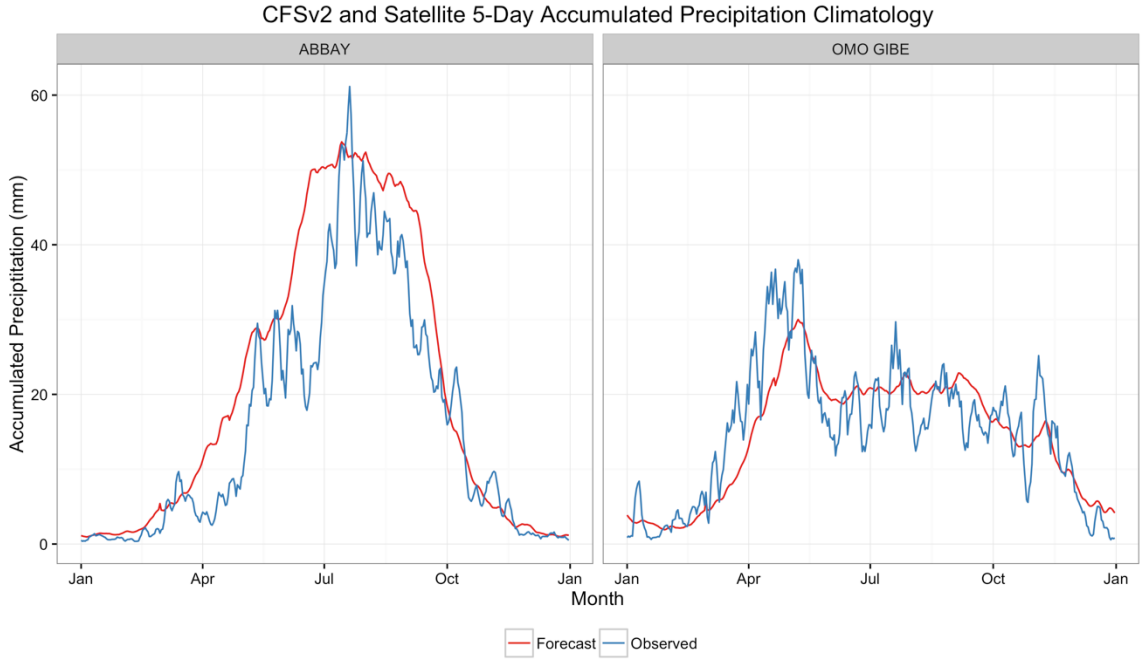

Figure 6-3 Daily Climatology of 5-day accumulated precipitation for observed (blue) and CFSv2 forecasts (red).

The CFSv2 also captures the magnitude of precipitation, corresponding well to the CHIRPS precipitation.

An initial assessment of forecast skill was provided by correlating the raw CFSv2 forecasts with CHIRPS precipitation. Figure 6-4 shows the monthly average correlations for a 5-day accumulation period, and forecasts at lead times of 0 to 60 days.

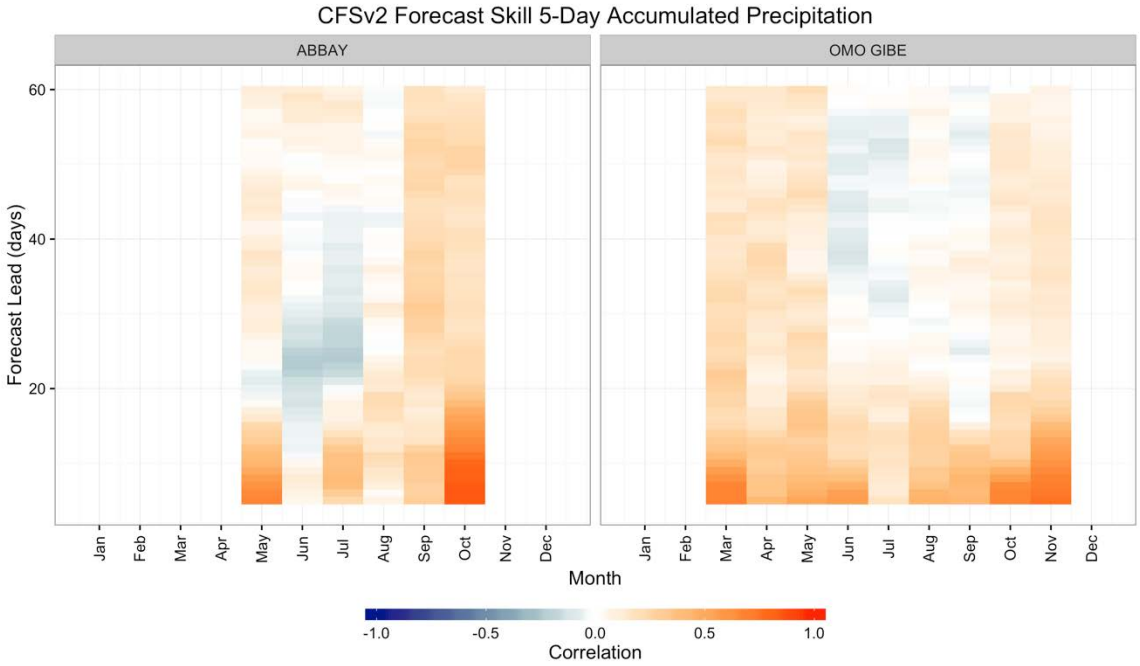

Figure 6-4 Correlation between 5-day accumulated CFSv2 forecasts and 5-day accumulated observed precipitation for forecast lead-times of 0-60 days. Months with less than 5% of the annual total precipitation have been masked out.

In the Abbay basin, correlations are highest in the shoulder months, May and October, at short lead-times, and remain strong for approximately 10 days. Correlations even at short leads during the peak season of June-September are marginal. The Omo Gibe basin shows similar behavior where the March and November shoulder months have strong correlations at short lead times that weaken around 10 days, and poor correlation skill during the April-October peak season. These low correlations at longer lead times and during peak precipitation seasons suggest that the dynamical model does not well represent the large-scale features like MJO that influence precipitation at the S2S timescale during these periods. We first explore how bias-correction and QR using the ensemble mean, ensemble quantile timeseries, and persistence forecasts can improve raw forecast skill at longer lead-times at during peak precipitation seasons. We then show the how including a measure of MJO in QR can further increase skill.

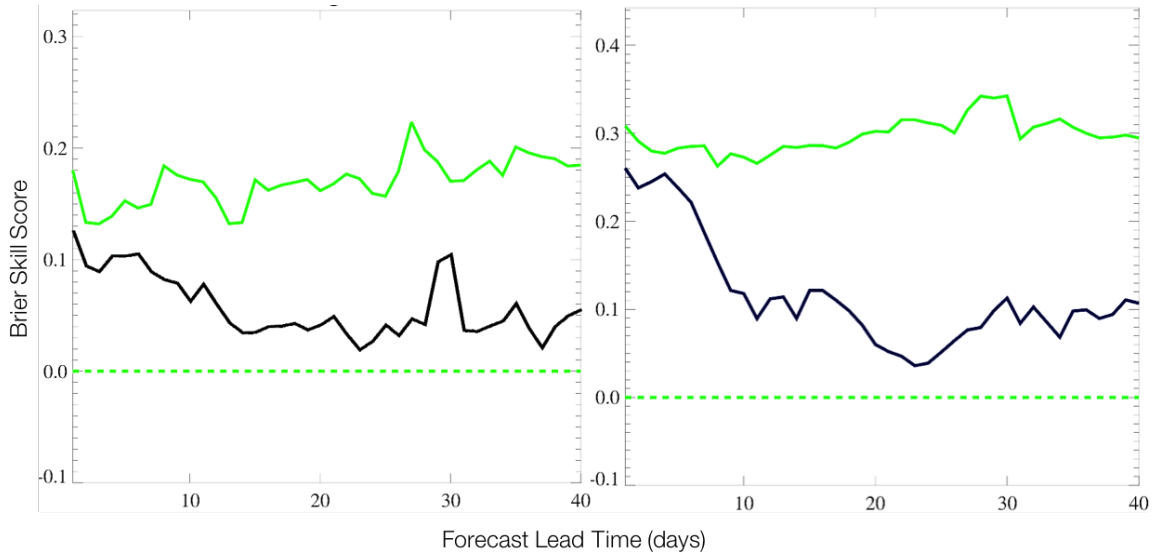

Figure 6-5 Brier Skill Scores (80th percentile exceedance) for the Abbay (left) and Omo Gibe (right) 5-day accumulated precipitation forecasts for the June-August Kiremt season using persistence as the reference forecasts. Green dashed line shows the skill of the persistence reference forecast, black line shows the CFSv2 QR without the naïve MJO forecast, and the green solid line shows the CFSv2 QR with the naïve MJO forecast.

The Briar Skill Score (BSS) is a metric to assess the forecast skill in capturing precipitation above a selected threshold. BSS evaluates a forecast relative to a reference forecast where the larger the score (up to a value of one, the better the forecast. A value of zero indicates no improvement over the reference forecast (Wilks 2001). Exceedance BSS indicate forecast skill in capturing precipitation above a selected threshold. For our evaluation, we used "persistence" as a reference forecast, assuming conditions at the initialization date will remain constant over the forecast period, and selected 80th percentile exceedance. 80th percentile exceedance represents larger precipitation events that could present natural hazard risks through flooding, and pose water management challenges. Figure 6-5 shows the 80th percentile exceedance BSS for the CFSv2 QR 5-day accumulated precipitation forecast for the June-August Kiremt season as compared to the persistence forecast (black line). The Abbay basin has a BSS above 0.05 for approximately 10 days while the Omo Gibe basin has a BSS above 0.05 for approximately 20 days. While marginal, this indicates that the bias-corrected and

post-processed forecasts have some skill above persistence. When the MJO is used to condition the forecasts in post-processing (Figure 6-5, green line) additional skill over persistence is evident. Forecast skill as measured by BSS remains consistent at lead times from 0 days out to 40 days for both the Abbay and Omo Gibe basins. At a 0-day lead time the Abbay basin has a BSS of 0.18 and the Omo Gibe a BSS of 0.3 and the BSS remains the same out to a 40-day lead time. We attribute this consistent skill to the predictable behavior of the MJO once active and its consistent and predictable influence on precipitation anomalies during the Kiremt season (Chapter 4).

5.4 DISCUSSION

Existing dynamical-model based precipitation forecasts at the S2S timescale suffer from a lack of skill during peak precipitation periods for watersheds in Ethiopia. This includes the June-August Kiremt season which is the peak precipitation season for the Abbay River basin and a major precipitation season for the Omo Gibe River basin. In both basins there has been recent and continued development of water management infrastructure including dams, hydropower production and water diversion for agricultural purposes. The S2S timescale, from two weeks to two months, represents a key decision-making period for water managers and rely heavily on skillful forecasts if precipitation. We have demonstrated through bias-correction and QR, significant improvements in forecast skill at these timescales can be obtained. Even increased improvements in forecast skill are seen when leveraging the predictable behavior of and predictable influence on precipitation of the MJO. While prior studies have shown that the MJO influences precipitation during the spring and fall in East Africa, this study is the first to also show it also influences Kiremt season precipitation. The technique for leveraging these influences in improving forecast skill is novel and represents a step towards improving the skill of S2S forecasts in this region. The naïve MJO forecast approach provides a parsimonious way of quantifying the MJO influence on

precipitation during the Kiremt season, and when used in QR demonstrates notable utility in improving S2S forecast skill. The QR approach is flexible and the approach described in this study clearly improved S2S forecasts of precipitation. Additional refinement of this approach including investigating methods for quantifying the influence of other dynamical features on precipitation like the El Nino Southern Oscillation and incorporating them into QR could show what additional improvements in skill are possible.

6 FINAL DISCUSSION

The four chapters presented examine how the strength of the West African and East African monsoons have varied both subseasonally and interseasonally over the 20th and early 21st centuries. While prominent events including a wet period in the middle of the 20th century and a drought in the late 20th century have been well documented, this research highlights how this interseasonal variability can be linked to changes in large-scale teleconnections. In examining these teleconnections this research also notes that these teleconnections have entered into a new epoch. In West Africa, local features, namely the Gulf of Guinea and to a lesser extent the Tropical Atlantic, have strengthened their connections with monsoon precipitation while more remote teleconnections, the Indian Ocean, and ENSO, have weakened. In East Africa, teleconnections with the tropical Atlantic, the tropical Indian Ocean, and the tropical Pacific have all weakened since 2000. While the dynamical shifts associated with this new teleconnection behavior have not yet fully been investigated, this understanding can be used to inform the development of tools used by decision makers, including forecasting models, where assumptions about past behavior, statistical relationships, or the use of analog years should consider these changes. Changes to these teleconnections and precipitation in both West and East Africa can be explained in part by the fact that both regions are influenced by the same teleconnections and even share some dynamical features, including Tropical Easterly Jet (TEJ). In addition to the interseasonal variability, this research investigated subseasonal behavior in the two monsoon regions, examining the early, middle, and late subseasons. Anomaly patterns during these three subseasons is not consistent, with a much more mixed signal during the early season, and more consistent wet or dry signals during the middle and late seasons. This behavior was noted in both West and East Africa and has also be documented with the Indian Monsoon. While the dynamical explanations for this behavior were not investigated directly, it is suspected that in all three regions, sea-surface teleconnections that largely influence monsoon strength and precipitation variability have much less of an influence relative to

local processes in the early subseason, during the establishment of the monsoon. Once the monsoon has been established, these teleconnections begin to dominate and explain the stronger and more homogenous anomaly patterns seen in precipitation during the middle and late subseasons. A final contribution from examining monsoon dynamics is showing the influence of the Madden-Julian Oscillation (MJO) on Kiremt season (June-September) precipitation in East Africa. The MJO has been shown to influence spring and fall season precipitation, but this is first study that shows this influence also exists during the summer, in northern East Africa, where the East African Monsoon is active. Through two case studies this examination of monsoon dynamics was extended to developing locally-targeted information, highlighting the development of tools and methods.

Developing regions stand to benefit immensely from improved climate-based information. Often existing information including weather forecasts and climate projections are not provided at spatial and temporal scales relevant for decision-making. The West and East African case studies highlight two approaches, stochastic weather generators, and post-processing numerical weather prediction forecasts, to address these limitations and produce actionable information at scales relevant for addressing local issues. Weather generators have used in a range of applications, however they have seen limited use by the livestock community. Especially in a region like West Africa where livestock herding is a major livelihood, and there are limited means for coping with the impacts of climate change, information about risk and forecasts can be immensely useful. The weather generator developed was modified to specifically allow the modeling of weather variables needed to quantify livestock heat stress and was developed to produce current-condition estimates, seasonal forecasts, and long-term projections. The seasonal forecasts were shown to have skill, and could offer benefits to the livestock herding community by allowing them to make better informed decisions on where and when to migrate. This forecast information could also be used by governmental agencies or others to develop forecast-based insurance programs which provide for a payout if conditions lead to a loss of livestock. These programs have been used with

success by the agricultural community in insuring against crop loss, and with a tool for forecasting livestock heat stress, could be developed for the livestock community as well.

Subseasonal to seasonal timescales (S2S; two weeks out to two months) have been a focus for developing improved forecasts. In this period, water managers are often evaluating possible operational decisions to meet monthly or seasonal requirements, however, often have to rely on forecasts with limited skill. Numerical weather prediction skill begins to fall in the S2S period, however can be improved when longer-term signals can be incorporated into the forecasts. In East Africa, the MJO influences precipitation variability when active, and this research shows how incorporating in the MJO into post-processing of NWP forecasts can improve predictability of precipitation in the S2S period.

The monsoon dynamics examined in this dissertation highlight several areas for future investigation including examining the dynamical causes of teleconnection variability over the last century and a better understanding of early subseason precipitation anomalies. The tools developed to produce local, issue-specific information have demonstrated skill in producing actionable information, however interaction with the user communities is needed to ensure that the information is relevant and disseminated in an effective way. These tools, the stochastic weather generator, and post-processing on NWP forecasts, have broad applicability, and while this research focused on specific case studies, it is hoped that by demonstrating their use in these case studies, along with their flexibility, other user communities will also see their utility.

7 REFERENCES

Andrés Ferreyra, R., G. P. Podestá, C. D. Messina, D. Letson, J. Dardanelli, E. Guevara, and S. Meira, 2001: A linked-modeling framework to estimate maize production risk associated with ENSO-related climate variability in Argentina. *Agric. For. Meteorol.*, **107**, 177–192, doi:10.1016/S0168-1923(00)00240-9.

Apipattanavis, S., F. Bert, G. Podestá, and B. Rajagopalan, 2010: Linking weather generators and crop models for assessment of climate forecast outcomes. *Agric. For. Meteorol.*, **150**, 166–174, doi:10.1016/j.agrformet.2009.09.012.

Bader, J., and M. Latif, 2003: The impact of decadal-scale Indian Ocean sea surface temperature anomalies on Sahelian rainfall and the North Atlantic Oscillation. *Geophys. Res. Lett.*, **30**, 1–4, doi:10.1029/2003GL018426.

Battisti, D., and R. Naylor, 2009: Historical warnings of future food insecurity with unprecedented seasonal heat. *Science (80-.).*, **323**, 240–244.

Beersma, J. J., and T. Adri Buishand, 2003: Multi-site simulation of daily precipitation and temperature conditional on the atmospheric circulation. *Clim. Res.*, **25**, 121–133, doi:10.3354/cr025121.

Benedict, J. J., M. S. Pritchard, and W. D. Collins, 2015: Sensitivity of MJO propagation to a robust positive Indian Ocean dipole event in the superparameterized CAM. *J. Adv. Model. Earth Syst.*, **7**, 1901–1917, doi:10.1002/2015MS000530.

Berhane, F., and B. Zaitchik, 2014: Modulation of Daily Precipitation over East Africa by the Madden–Julian Oscillation. *J. Clim.*, **27**, 6016–6034, doi:10.1175/JCLI-D-13-00693.1.

Bjørnar Bremnes, J., 2004: Probabilistic Forecasts of Precipitation in Terms of Quantiles Using NWP Model Output. *Mon. Weather Rev.*, **132**, 338–347, doi:10.1175/1520-0493(2004)132<0338:PFOPIT>2.0.CO;2.

Block, P., and B. Rajagopalan, 2007: Interannual Variability and Ensemble Forecast of Upper Blue Nile Basin Kiremt Season Precipitation. *J. Hydrometeorol.*, **8**, 327–343, doi:10.1175/JHM580.1.

Brandsma, T., and T. A. Buishand, 1998: Simulation of extreme precipitation in the Rhine basin by nearest-neighbour resampling. *Hydrol. Earth Syst. Sci.*, **2**, 195–209, doi:10.5194/hess-2-195-1998.

Broman, D. P., B. Rajagopalan, and T. Hopson, Spatial and Temporal Variability of Sahelian West African Summer Monsoon Precipitation and Large-Scale Teleconnections. *J. Geophys. Res. Atmos.*,.

——, ——, and ——, Spatial and Temporal Variability of East African Kiremt Season Precipitation and Large-Scale Teleconnections. *Int. J. Climatol.*,.

Brown-Brandl, T. M., R. a. Eigenberg, and J. a. Nienaber, 2006: Heat stress risk factors of feedlot heifers. *Livest. Sci.*, **105**, 57–68, doi:10.1016/j.livsci.2006.04.025.

Buishand, T. A., and T. Brandsma, 2001: Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling. *Water Resour. Res.*, **37**, 2761–2776, doi:10.1029/2001WR000291.

Caniaux, G., H. Giordani, J.-L. L. Redelsperger, F. F. Guichard, E. Key, and M. Wade, 2011: Coupling between the Atlantic cold tongue and the West African monsoon in boreal spring and summer. *J. Geophys. Res.*, **116**, C04003, doi:10.1029/2010JC006570.

Caraway, N. M., J. L. McCreight, and B. Rajagopalan, 2014: Multisite stochastic weather generation using cluster analysis and k-nearest neighbor time series resampling. *J. Hydrol.*, **508**, 197–213, doi:10.1016/j.jhydrol.2013.10.054.

Caron, A., R. Leconte, and F. Brissette, 2008: An Improved Stochastic Weather Generator for Hydrological Impact Studies. *Can. Water Resour. J.*, **33**, 233–256.

Charney, J. G., 1975: Dynamics of deserts and drought in the Sahel. *Q. J. R. Meteorol. Soc.*, **101**, 193–202.

Chen, T.-C., and H. van Loon, 1987: Interannual Variation of the Tropical Easterly Jet. *Mon. Weather Rev.*, **115**, 1739–1759, doi:10.1175/1520-0493(1987)115<1739:IVOTTE>2.0.CO;2.

Dai, A., P. J. Lamb, K. E. Trenberth, M. Hulme, P. D. Jones, and P. Xie, 2004: The recent Sahel drought is real. *Int. J. Climatol.*, **24**, 1323–1331, doi:10.1002/joc.1083.

Dieppois, B., A. Durand, M. Fournier, A. Diedhiou, B. Fontaine, N. Massei, Z. Nouaceur, and D. Sebag, 2014: Low-frequency variability and zonal contrast in Sahel rainfall and Atlantic sea surface temperature teleconnections during the last century. *Theor. Appl. Climatol.*, **121**, 139–155, doi:10.1007/s00704-014-1229-5.

Diro, G. T., D. I. F. Grimes, and E. Black, 2011: Teleconnections between Ethiopian summer rainfall and sea surface temperature: Part I-observation and modelling. *Clim. Dyn.*, **37**, 103–119, doi:10.1007/s00382-010-0896-x.

Dunn, R. J. H., K. M. Willett, P. W. Thorne, E. V Woolley, I. Durre, A. Dai, D. E. Parker, and R. S. Vose, 2012: HadISD : A quality controlled global synoptic report database for selected variables at long-term stations from 1973-2011.

Easterling, W., and M. Apps, 2005: Assesing the Consequences of climate Change for Food and Foret Resources: A View from the IPCC. *Clim. Change*, **70**, 165–189.

Eigenberg, R. a., T. M. Brown-Brandl, J. a. Nienaber, G. L. Hahn, R. a. Eigenberg, J. a. Nienaber, and G. L. Hahn, 2005: Dynamic Response Indicators of Heat Stress in Shaded and Non-shaded Feedlot Cattle, Part 2: Predictive Relationships. *Biosyst. Eng.*, **90**, 451–462, doi:10.1016/j.biosystemseng.2005.02.001.

Eigenberg, R. a., T. M. Brown-Brandl, and J. a. Nienaber, 2007: Development of a livestock weather safety monitor for feedlot cattle. *Appl. Eng. Agric.*, **23**, 657–660.

Enfield, D. B., A. M. Mestas-Nunez, D. A. Mayer, and L. Cid-Cerrano, 1999: How ubiquitous is the dipole relationship in tropical Atlantic sea surface temperatures? *J. Geophys. Res.*, **104**, 7841–7848.

Erkyihun, S. T., B. Rajagopalan, E. Zagona, U. Lall, and K. Nowak, 2016: Wavelet-based time series bootstrap model for multidecadal streamflow simulation using climate indicators. *Water Resour. Res.*, **52**, 4061–4077, doi:10.1002/2016WR018696.

Field, C. B., and Coauthors, eds., 2014: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, New York, 668 pp.

Finch, V., 1986: Body Temperature in Beef Cattle: Its Control and Relevance to Production in the Tropics. *J. Anim. Sci.*, 531–542.

Flohn, H., 1987: Rainfall teleconnections in northern and northeastern Africa. *Theor. Appl. Climatol.*, **38**, 191–197, doi:10.1007/BF00867412.

Folland, C. K., T. N. Palmer, and D. E. Parker, 1986: Sahel rainfall and worldwide sea temperatures 1901-1985. *Nature*, **320**, 602–607.

Fontaine, B., and S. Janicot, 1998: Sea Surface Temperature Fields Associated with West African Rainfall Anomaly Types. *J. Clim.*, **9**, 2935–2940.

——, and S. Louvet, 2006: Sudan-Sahel rainfall onset: Definition of an objective index, types of years, and experimental hindcasts. *J. Geophys. Res.*, **111**, D20103, doi:10.1029/2005JD007019.

Friend, A. D., and Coauthors, 2014: Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 3280–3285, doi:10.1073/pnas.1222477110.

Funk, C., and Coauthors, 2015: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci. Data*, **2**, 150066, doi:10.1038/sdata.2015.66.

Furrer, E., and R. Katz, 2007: Generalized linear modeling approach to stochastic weather generators. *Clim. Res.*, **34**, 129–144.

Furrer, E. M., and R. W. Katz, 2008: Improving the simulation of extreme precipitation events by stochastic weather generators. *Water Resour. Res.*, **44**, n/a-n/a, doi:10.1029/2008WR007316.

Gaughan, J. B., 2006: Modelling the Components of Livestock Stress for Precision Animal Management. *2006 Portland, Oregon, July 9-12, 2006*, doi:10.13031/2013.21185.

——, T. L. Mader, S. M. Holt, and a Lisle, 2008: A new heat load index for feedlot cattle. *J. Anim. Sci.*, **86**, 226–234, doi:10.2527/jas.2007-0305.

Giannini, A., R. Saravannan, and P. Chang, 2003: Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. *Science (80-.).*, **302**, 1027–1030, doi:10.1126/science.1089357.

Gill, E. C., B. Rajagopalan, and P. Molnar, 2015a: Subseasonal variations in spatial signatures of ENSO on the Indian summer monsoon from 1901 to 2009. *J. Geophys. Res. Atmos.*, **120**, 8165–8185, doi:10.1002/2015JD023184.

Gill, E. C., B. Rajagopalan, and P. Molnar, 2015b: Subseasonal variations in spatial signatures of ENSO on the Indian summer monsoon from 1901 to 2009. *J. Geophys. Res. Atmos.*, **120**, 8165–8185, doi:10.1002/2015JD023184.

Gissila, T., E. Black, D. I. F. Grimes, and J. M. Slingo, 2004: Seasonal forecasting of the Ethiopian summer rains. *Int. J. Climatol.*, **24**, 1345–1358, doi:10.1002/joc.1078.

Greatrex, H., and Coauthors, 2015: *Scaling up index insurance for smallholder farmers : Recent evidence and insights*.

Grondona, M. O., G. P. Podesta, M. Bidegain, M. Marino, and H. Hordij, 2000: A stochastic precipitation generator conditioned on ENSO phase: A case study in southeastern South America. *J. Clim.*, **13**, 2973–2986, doi:10.1175/1520-0442(2000)013<2973:ASPGCO>2.0.CO;2.

Hagos, S. M., and K. H. Cook, 2008: Ocean warming and late-twentieth-century Sahel drought and recovery. *J. Clim.*, **21**, 3797–3814, doi:10.1175/2008JCLI2055.1.

Hahn, G. L., J. B. Gaughan, T. L. Mader, and R. A. Eigenberg, 2009: Thermal Indices and Their Applications for Livestock Environments. *Livestock Energetics and Thermal Environmental Management*, 113–130.

Hansen, J. W., and T. Mavromatis, 2001: Correcting low-frequency variability bias in stochastic weather generators. *Agric. For. Meteorol.*, **109**, 297–310, doi:10.1016/S0168-1923(01)00271-4.

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister, 2014: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *Int. J. Climatol.*, **34**, 623–642, doi:10.1002/joc.3711.

Harrold, T. I., 2003: A nonparametric model for stochastic generation of daily rainfall amounts. *Water Resour. Res.*, **39**, 1343, doi:10.1029/2003WR002570.

Hopson, T. M., and P. J. Webster, 2010a: A 1–10-Day Ensemble Forecasting Scheme for the Major River Basins of Bangladesh: Forecasting Severe Floods of 2003–07*. *J. Hydrometeorol.*, **11**, 618–641, doi:10.1175/2009JHM1006.1.

Hopson, T. M., and P. J. Webster, 2010b: A 1–10-Day Ensemble Forecasting Scheme for the Major River Basins of Bangladesh: Forecasting Severe Floods of 2003–07*. *J. Hydrometeorol.*, **11**, 618–641, doi:10.1175/2009JHM1006.1.

Hulme, M., and N. Tosdevin, 1989: The Tropical easterly Jet and Sudan rainfall: A review. *Theor. Appl. Climatol.*, **39**, 179–187, doi:10.1007/BF00867945.

Hurrell, J. W., J. J. Hack, D. Shea, J. M. Caron, and J. Rosinski, 2008: A new sea surface temperature and sea ice boundary dataset for the community atmosphere model. *J. Clim.*, **21**, 5145–5153, doi:10.1175/2008JCLI2292.1.

Jackson, B., S. E. Nicholson, and D. Klotter, 2009: Mesoscale Convective Systems over Western Equatorial Africa and Their Relationship to Large-Scale Circulation. *Mon. Weather Rev.*, **137**, 1272–1294, doi:10.1175/2008MWR2525.1.

Janicot, S., F. Mounier, N. M. J. Hall, S. Leroux, B. Sultan, and G. N. Kiladis, 2009:

Dynamics of the West African monsoon. Part IV: Analysis of 25-90-day variability of convection and the role of the Indian monsoon. *J. Clim.*, **22**, 1541–1565, doi:10.1175/2008JCLI2314.1.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Am. Meteorol. Soc.*, doi:http://dx.doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Katz, R., 2002: Statistics of Extremes in Climatology and Hydrology. *Adv. Water Resour.*, **25**, 1287–1304.

Katz, R. W., 1977: Precipitation as a Chain-Dependent Process. *J. Appl. Meteorol.*, **16**, 671–676, doi:10.1175/1520-0450(1977)016<0671:PAACDP>2.0.CO;2.

Kilsby, C. G., and Coauthors, 2007: A daily weather generator for use in climate change studies. *Environ. Model. Softw.*, **22**, 1705–1719, doi:10.1016/j.envsoft.2007.02.005.

Kim, Y., R. Katz, B. Rajagopalan, G. Podestá, and E. Furrer, 2012: Reducing overdispersion in stochastic weather generators using a generalized linear modeling approach. *Clim. Res.*, **53**, 13–24, doi:10.3354/cr01071.

Kirtman, B. P., and Coauthors, 2014: The North American multimodel ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bull. Am. Meteorol. Soc.*, **95**, 585–601, doi:10.1175/BAMS-D-12-00050.1.

Kleiber, W., R. W. Katz, and B. Rajagopalan, 2012: Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. *Water Resour. Res.*, **48**, 1–17, doi:10.1029/2011WR011105.

——, ——, and ——, 2013: Daily minimum and maximum temperature simulation over complex terrain. *Ann. Appl. Stat.*, **7**, 588–612, doi:10.1214/12-AOAS602.

Knievel, J. C., and Coauthors, 2017: Mesoscale ensemble weather prediction at U.S. Army Dugway Proving Ground, Utah. *Weather Forecast.*, **32**, 2195–2217, doi:10.1175/WAF-D-17-0049.1.

Korecha, D., and A. G. Barnston, 2007: Predictability of June–September Rainfall in Ethiopia. *Mon. Weather Rev.*, **135**, 628–650, doi:10.1175/MWR3304.1.

Koster, R. D., and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **305**, 1138–1140, doi:10.1126/science.1100217.

Krishnaswamy, J., S. Vaidyanathan, B. Rajagopalan, M. Bonell, M. Sankaran, R. S. Bhalla, and S. Badiger, 2015: Non-stationary and non-linear influence of ENSO and Indian Ocean Dipole on the variability of Indian monsoon rainfall and extreme rain events. *Clim. Dyn.*, **45**, 175–184, doi:10.1007/s00382-014-2288-0.

Kug, J. S., F. F. Jin, and S. Il An, 2009: Two types of El Niño events: Cold tongue El Niño and warm pool El Niño. *J. Clim.*, **22**, 1499–1515, doi:10.1175/2008JCLI2624.1.

Kumar, K., B. Rajagopalan, and M. Cane, 1999: On the weakening relationship between the indian monsoon and ENSO. *Science*, **284**, 2156–2159, doi:DOI 10.1126/science.284.5423.2156.

Lebel, T., and A. Ali, 2009: Recent trends in the Central and Western Sahel rainfall regime (1990–2007). *J. Hydrol.*, **375**, 52–64, doi:10.1016/j.jhydrol.2008.11.030.

——, A. Diedhiou, and H. Laurent, 2003: Seasonal cycle and interannual variability of the Sahelian rainfall at hydrological scales. *J. Geophys. Res. Atmos.*, **108**, 1–11, doi:10.1029/2001JD001580.

Losada, T., B. Rodriguez-Fonseca, E. Mohino, J. Bader, S. Janicot, and C. R. Mechoso, 2012: Tropical SST and Sahel rainfall: A non-stationary relationship. *Geophys. Res. Lett.*, **39**, 1–7, doi:10.1029/2012GL052423.

Lu, J., and D. T.L., 2005: Ocean forcing of the late 20th century Sahel drought. *Geophys. Res. Lett.*,.

Mader, T. L., L. J. Johnson, and J. B. Gaughan, 2010: A comprehensive index for assessing environmental stress in animals. *J. Anim. Sci.*, **88**, 2153–2165, doi:10.2527/jas.2009-2586.

Mahoney, W. P., and Coauthors, 2012: A Wind Power Forecasting System to Optimize Grid Integration. *Sustain. Energy, IEEE Trans.*, **3**, 670–682, doi:10.1109/TSTE.2012.2201758.

Maity, R., and D. Nagesh Kumar, 2006: Bayesian dynamic modeling for monthly Indian summer monsoon rainfall using El Ni??o-Southern Oscillation (ENSO) and Equatorial Indian Ocean Oscillation (EQUINOO). *J. Geophys. Res. Atmos.*, **111**, 1–12, doi:10.1029/2005JD006539.

McCullagh, P., and J. A. Nelder, 1989: Generalized Linear Models, no. 37 in Monograph on Statistics and Applied Probability.

Meza, F. J., 2005: Variability of reference evapotranspiration and water demands. Association to ENSO in the Maipo river basin, Chile. *Glob. Planet. Change*, **47**, 212–220, doi:10.1016/j.gloplacha.2004.10.013.

Mohr, K. I., and C. D. Thorncroft, 2006: Intense convective systems in West Africa and their relationship to the African easterly jet. *Q. J. R. Meteorol. Soc.*, **132**, 163–176, doi:10.1256/qj.05.55.

Monteith, J., and M. Unsworth, 2014: *Principles of Environmental Physics*. Fourth. Elsevier Ltd,.

Moran, V., and M. N. Ward, 1998: ENSO teleconnections with climate variability in the European and African sectors. *Weather*, **53**, 287–295.

Nardone, A., B. Ronchi, N. Lacetera, M. S. S. Ranieri, and U. Bernabucci, 2010: Effects of climate changes on animal production and sustainability of livestock systems. *Livest. Sci.*, **130**, 57–69, doi:10.1016/j.livsci.2010.02.011.

Nesbitt, S. W., E. J. Zipser, and Zipser, 2003: The diurnal cycle of rainfall and convective intensity according to three years of TRMM measurements. *J. Clim.*, **16**, 1456–1475.

Nicholson, S. E., 2000: Land Surface Processes and Sahel Climate. *Rev. Geophys.*, **38**, 117–139.

——, 2005: On the question of the "recovery" of the rains in the West African Sahel. *J. Arid Environ.*, **63**, 615–641, doi:10.1016/j.jaridenv.2005.03.004.

——, 2013: The West African Sahel: A review of recent studies on the rainfall regime and its interannual variability. *ISRN Meteorol.*, **2013**, 32, doi:10.1155/2013/453521.

——, 2014: The predictability of rainfall over the Greater Horn of Africa. Part I. Prediction of seasonal rainfall. *J. Hydrometeorol.*, **15**, 140117143344004, doi:10.1175/JHM-D-13-062.1.

——, and J. Kim, 1997: The Relationshop of the El Nino-Southern Oscillation to African Rainfall. *Int. J. Climatol.*, **17**, 117–135, doi:10.1002/(SICI)1097-0088(199702)17:2<117::AID-JOC84>3.0.CO;2-O.

——, and J. P. P. Grist, 2001: A conceptual model for understanding rainfall variability in the West African Sahel on interannual and interdecadal timescales. *Int. J. Climatol.*, **21**, 1733–1757, doi:10.1002/joc.648.

——, C. J. C. Tucker, and M. B. Ba, 1998: Desertification, drought, and surface vegetation: an example from the West African Sahel. *Bull. Am. Meteorol. Soc.*, **79**, 815–829.

NOAA, 1976: *Livestock hot weather stress*. Kansas City, MO,.

Palmer, T. N., 1986: Influence of the Atlantic, Pacific, and Indian Oceans on Sahel rainfall. *Nature*, **322**, 251–253.

Parsons, D. J., a. C. Armstrong, J. R. Turnpenny, a. M. Matthews, K. Cooper, and J. a. Clark, 2001a: Integrated models of livestock systems for climate change studies. 2. Intensive systems. *Glob. Chang. Biol.*, **7**, 163–170, doi:10.1046/j.1365-2486.2001.00392.x.

——, A. C. Armstrong, J. R. Turnpenny, A. M. Matthews, K. Cooper, and J. A. Clark, 2001b: Integrated models of livestock systems for climate change studies. 1. Grazing systems. *Glob. Chang. Biol.*, **7**, 93–112, doi:10.1046/j.1365-2486.2001.00392.x.

Petris, G., 2009: dlm : an R package for Bayesian analysis of Dynamic Linear Models. *Univ. Arkansas*, 1–23.

Podestá, G., and Coauthors, 2009: Decadal climate variability in the Argentine Pampas: regional impacts of plausible climate scenarios on agricultural systems. *Clim. Res.*, **40**, 199–210, doi:10.3354/cr00807.

Qian, B., S. Gameda, R. De Jong, P. Falloon, and J. Gornall, 2010: Comparing scenarios of canadian daily climate extremes derived using a weather generator. *Clim. Res.*, **41**, 131–149, doi:10.3354/cr00845.

Rajagopalan, B., and U. Lall, 1999: A k-nearest-neighbor Simulator for Daily Precipitation and Other Variables. *Water Resour. Res.*, **35**, 3089–3101, doi:10.1029/1999WR900028.

Rajagopalan, B., U. Lall, D. G. Tarboton, and D. S. Bowles, 1997: Multivariate nonparametric resampling scheme for generation of daily weather variables. *Stoch. Hydrol. Hydraul.*, **11**, 65–93, doi:10.1007/BF02428426.

Richardson, C. W., 1981: Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resour. Res.*, **17**, 182, doi:10.1029/WR017i001p00182.

Richardson, C. W., and D. a Wright, 1984: WGEN : A Model for Generating Daily Weather Variables By. *Agriculture*, **ARS-8**, 83.

Robertson, A. W., A. Kumar, M. Peña, and F. Vitart, 2015: Improving and Promoting Subseasonal to Seasonal Prediction. *Bull. Am. Meteorol. Soc.*, **96**, ES49-ES53, doi:10.1175/BAMS-D-14-00139.1.

Rosenzweig, C., and Coauthors, 2013: The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. *Agric. For. Meteorol.*, **170**, 166–182, doi:10.1016/j.agrformet.2012.09.011.

Saha, S., and Coauthors, 2014: The NCEP climate forecast system version 2. *J. Clim.*, **27**, 2185–2208, doi:10.1175/JCLI-D-12-00823.1.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. *Nature*, **401**, 360–363, doi:10.1038/43854.

Sankarasubramanian, A., and U. Lall, 2003: Flood quantiles in a changing climate: Seasonal forecasts and causal relations. *Water Resour. Res.*, **39**, 1–12, doi:10.1029/2002WR001593.

——, ——, N. Devineni, and S. Espinueva, 2009: The role of monthly updated climate forecasts in improving intraseasonal water allocation. *J. Appl. Meteorol. Climatol.*, **48**, 1464–1482, doi:10.1175/2009JAMC2122.1.

Scott, a J., and M. Knott, 1974: A Cluster Analysis Method for Grouping Means in the Analysis of Variance. *Biometrics*, **30**, 507–512, doi:10.2307/2529204.

Segele, Z. T., and P. J. Lamb, 2005: Characterization and variability of Kiremt rainy season over Ethiopia. *Meteorol. Atmos. Phys.*, **89**, 153–180, doi:10.1007/s00703-005-0127-x.

Segele, Z. T., P. J. Lamb, and L. M. Leslie, 2009a: Large-scale atmospheric circulation and global sea surface temperature associations with Horn of Africa June-September rainfall. *Int. J. Climatol.*, **29**, 1075–1100, doi:10.1002/joc.1751.

——, ——, and ——, 2009b: Seasonal-to-interannual variability of Ethiopia/horn of Africa monsoon. Part I: Associations of wavelet-filtered large-scale atmospheric circulation and global sea surface temperature. *J. Clim.*, **22**, 3396–3421, doi:10.1175/2008JCLI2859.1.

Seleshi, Y., and U. Zanke, 2004: Recent changes in rainfall and rainy days in Ethiopia. *Int. J. Climatol.*, **24**, 973–983, doi:10.1002/joc.1052.

Sharif, M., and D. H. Burn, 2007: Improved K -Nearest Neighbor Weather Generating Model. *J. Hydrol. Eng.*, **12**, 42–51, doi:10.1061/(ASCE)1084-0699(2007)12:1(42).

Shimizu, M. H., and T. Ambrizzi, 2016: MJO influence on ENSO effects in precipitation and temperature over South America. *Theor. Appl. Climatol.*, **124**, 291–301, doi:10.1007/s00704-015-1421-2.

Shinoda, T., and W. Han, 2005: Influence of the Indian Ocean Dipole on Atmospheric Subseasonal Variability. *J. Clim.*, **18**, 3891–3909, doi:10.1175/JCLI3510.1.

Shukla, S., J. Roberts, A. Hoell, C. C. Funk, F. Robertson, and B. Kirtman, 2016: Assessing North American multimodel ensemble (NMME) seasonal forecast skill to assist in the early warning of anomalous hydrometeorological events over East Africa. *Clim. Dyn.*, 1–17, doi:10.1007/s00382-016-3296-z.

Silverman, B. W., 1986: *Density estimation for statistics and data analysis*. CRC press,.

Stern, R. D., and R. Coe, 1984: A Model Fitting Analysis of Daily Rainfall Data. *J. R. Stat. Soc. Ser. A*, **147**, 1–34, doi:10.2307/2981736.

Stocker, T. F., and Coauthors, eds., 2013: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovern- mental Panel on Climate Change*. Cambridge University Press, New York, 1535 pp.

Surminski, S., L. M. Bouwer, and J. Linnerooth-bayer, 2016: How insurance can support climate resilience. *Nat. Clim. Chang. Gr.*, **6**, 333–334, doi:10.1038/nclimate2979.

Thom, E. C., 1959: The Discomfort Index. *Weatherwise*, **12**, 57–61, doi:10.1080/00431672.1959.9926960.

Thompson, V. a., and Coauthors, 2013: The development of a dynamic, mechanistic, thermal balance model for Bos indicus and Bos taurus. *J. Agric. Sci.*, **152**, 464–482, doi:10.1017/S0021859613000506.

Thornton, P., and Coauthors, 2006: *Mapping climate vulnerability and poverty in Africa*. Nairobi, Kenya, 171 pp.

Thornton, P. K. K., J. van de Steeg, a. Notenbaert, and M. Herrero, 2009: The impacts of climate change on livestock and livestock systems in developing countries: A review of what we know and what we need to know. *Agric. Syst.*, **101**, 113–127, doi:10.1016/J.Agsy.2009.05.002.

Torrence, C., and P. J. Webster, 1998: The annual cycle of persistence in the El Niño/Southern Oscillation. *Q. J. R. Meteorol. Soc.*, **124**, 1985–2004, doi:10.1002/qj.49712455010.

Towler, E., B. Rajagopalan, E. Gilleland, R. S. Summers, D. Yates, and R. W. Katz, 2010: Modeling hydrologic and water quality extremes in a changing climate: A statistical approach based on extreme value theory. *Water Resour. Res.*, **46**, n/a--n/a, doi:10.1029/2009WR008876.

Trenberth, K. E., 1997: The Definition of El Nino. *Bull. Am. Meteorol. Soc.*, **78**, 2771–2777, doi:10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2.

United Nations General Assembly, 2015: *Transforming our world: the 2030 Agenda for Sustainable Development*.

Verdin, A., B. Rajagopalan, W. Kleiber, and R. W. Katz, 2014: Coupled stochastic weather generation using spatial and generalized linear models. *Stoch. Environ. Res. Risk Assess.*, **29**, 347–356, doi:10.1007/s00477-014-0911-6.

——, ——, ——, G. Podestá, and F. Bert, 2018: A conditional stochastic weather generator for seasonal to multi-decadal simulations. *J. Hydrol.*, **556**, 835–846, doi:10.1016/j.jhydrol.2015.12.036.

Vizy, E. K., and K. H. Cook, 2002: Development and application of a mesoscale climate model for the tropics : Influence of sea surface temperature anomalies on the West African monsoon. **107**.

Webster, P. J., 2013: Improve weather forecasts for the developing world. *Nature*, **493**, 17–19, doi:10.1038/493017a.

Webster, P. J., and Coauthors, 2010: Extended-Range Probabilistic Forecasts of Ganges and Brahmaputra Floods in Bangladesh. *Bull. Am. Meteorol. Soc.*, **91**, 1493–1514, doi:10.1175/2010BAMS2911.1.

West, M., and J. Harrison, 1997: *Bayesian Forecasting and Dynamic Models*. 2nd ed. Springer, Berlin,.

Wheeler, M. C., and H. H. Hendon, 2004: An All-Season Real-Time Multivariate MJO Index: Development of an Index for Monitoring and Prediction. *Mon. Weather Rev.*, **132**, 1917–1932, doi:http://dx.doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

Wilby, R. ., C. . Dawson, and E. . Barrow, 2002: sdsm — a decision support tool for the assessment of regional climate change impacts. *Environ. Model. Softw.*, **17**,

145–157, doi:10.1016/S1364-8152(01)00060-3.

Wilks, D. S., 2001: A skill score based on economic value for probability forecasts. *Meteorol. Appl.*, **8**, 209–219, doi:10.1017/S1350482701002092.

Wilks, D. S., 2008: High-resolution spatial interpolation of weather generator parameters using local weighted regressions. *Agric. For. Meteorol.*, **148**, 111–120, doi:10.1016/j.agrformet.2007.09.005.

Williams, A. P., and Coauthors, 2012: Recent summer precipitation trends in the Greater Horn of Africa and the emerging role of Indian Ocean sea surface temperature. *Clim. Dyn.*, **39**, 2307–2328, doi:10.1007/s00382-011-1222-y.

Wilson, E. A., A. L. Gordon, and D. Kim, 2013: Observations of the madden julian oscillation during Indian ocean dipole events. *J. Geophys. Res. Atmos.*, **118**, 2588–2599, doi:10.1002/jgrd.50241.

Xue, Y., and Coauthors, 2016: West African monsoon decadal variability and surface-related forcings: second West African Monsoon Modeling and Evaluation Project Experiment (WAMME II). *Clim. Dyn.*, **47**, 3517–3545, doi:10.1007/s00382-016-3224-2.

Yang, C., R. E. Chandler, V. S. Isham, and H. S. Wheater, 2005: Spatial-temporal rainfall simulation using generalized linear models. *Water Resour. Res.*, **41**, 1–13, doi:10.1029/2004WR003739.

Yanto, B. Rajagopalan, and E. Zagona, 2016: Space–time variability of Indonesian rainfall at inter-annual and multi-decadal time scales. *Clim. Dyn.*, doi:10.1007/s00382-016-3008-8.

Yates, D., 2003: A technique for generating regional climate scenarios using a nearest-neighbor algorithm. *Water Resour. Res.*, **39**, 1–15, doi:10.1029/2002WR001769.

Zaitchik, B. F., 2017: Madden-Julian Oscillation impacts on tropical African precipitation. *Atmos. Res.*, **184**, 88–102, doi:10.1016/j.atmosres.2016.10.002.

8 APPENDIX - STOCHASTIC WEATHER GENERATOR DETAIL

Equations describing the construction of the stochastic weather generator without covariates are shown below.

8.1 STOCHASTIC WEATHER GENERATOR WITHOUT COVARIATES

8.1.1 PRECIPITATION OCCURRENCE

Precipitation occurrence is assumed to have a binomial (or Bernoulli) distribution with a no rain day represented by Jt = 0 and a rain day represented by $J_t = 1$. A binomial link function is used and precipitation occurrence is conditioned on the occurrence or non-occurrence of precipitation the previous day.

$$\ln\left(\frac{p_o}{1-p_o}\right) = \mu + \alpha J_{t-1} + \beta_{o,1} C_t + \beta_{o,2} S_t + \gamma_1 C_t J_{t-1} + \gamma_2 S_t J_{t-1}$$

Above C_t and S_t are sine and cosine harmonic terms, calculated as $\cos(R \times 2\pi t/n)$and $\sin(R \times 2\pi t/n)$ respectively. t is the number of days since January 1, n is the number of days om the year, and R is an integer. R values of 1-5 are used to capture the seasonal behavior. The model returns a value between 0 and 1, which is then transformed, to binary 0,1 using:

$$\mathrm{f}(x)\begin{cases}1, \text{if } \dfrac{\exp p_o}{1+\exp p_o} < \delta \\ 0, \quad \text{otherwise}\end{cases}$$

where δ is a random number between 0 and 1.

8.1.2 PRECIPITATION INTENSITY

Precipitation intensity is assumed to follow a Gamma distribution with an inverse link function.

$$\ln(p_i) = \mu + \beta_{i,1} C_t + \beta_{i,2} S_t$$

This model, as with occurrence, contains the cosine and sine harmonic terms.

8.1.3 TEMPERATURE

Temperature is modeled as a Gaussian with an identity link function. The previous day's minimum and maximum temperatures, along with precipitation occurrence, and harmonic terms are included in the model.

$$t_{n,t} = \mu_{n,0} + \mu_{n,1} + \gamma_{tn} t_{n,t-1} + \zeta_{tx} t_{x,t-1} + C_t + \beta_{tn,1} C_t + \beta_{tn,2} S_t + \eta_{t,x} P_{o,t} + \varepsilon_{tn}$$

$$t_{x,t} = \mu_{x,0} + \mu_{x,1} + \gamma_{tn} t_{x,t-1} + \zeta_{tx} t_{x,t-1} + C_t + \beta_{tx,1} C_t + \beta_{tx,2} S_t + \eta_{t,x} P_{o,t} + \varepsilon_{tx}$$

8.1.4 RELATIVE HUMIDITY

Let $J_{n,t}$ be the minimum relative humidity at time t, and $J_{x,t}$ be the maximum relative humidity at time with $0 \leq J_{n,t}, J_{x,t} \leq 1$. Relative humidity is modeled using a binomial link function and is conditioned on the value of relative humidity from the previous day, as well as the modeled minimum relative humidity for the current day in the case of maximum relative humidity.

$$\ln \frac{h_n}{1 - h_n} = \mu + \alpha_1 J_{n,t-1} + \alpha_2 J_{x,t-1} + \beta_{hn,1} C_t + \beta_{hn,2} S_t + \gamma_1 P_{0,t} + \zeta_1 t_{n,t-1} + \zeta_2 t_{x,t-1} + \eta_1 t_{n,t} + \eta_2 t_{x,t}$$

$$\ln \frac{h_x}{1 - h_x} = \mu + \alpha_1 J_{n,t-1} + \alpha_2 J_{x,t-1} + \beta_{hn,1} C_t + \beta_{hn,2} S_t + \gamma_1 P_{0,t} + \zeta_1 t_{n,t-1} + \zeta_2 t_{x,t-1} + \eta_1 t_{n,t} + \eta_2 t_{x,t} + \nu h_n$$

Relative humidity is also dependent on the previous day?s minimum and maximum temperature, the concurrent day minimum and maximum temperature, and the concurrent occurrence of precipitation. Taking the inverse logit of the modeled value and multiplying by 100 gives a value $0 \leq RH \leq 100$.

$$\mathrm{RH} = \exp(h_n, h_x)/\big(1 + \exp(h_n, h_x)\big) \times 100$$

8.1.5 WIND SPEED

Wind Speed is modeled as a Gaussian with an identity link function:

$$\mathrm{WS} = \mu + \beta_{i,1} C_t + \beta_{i,2} S_t + + \gamma_1 P_{o,t} + \zeta_1 t_{n,t-1} + \zeta_2 t_{x,t-1} + \nu_1 t_{n,t} + \nu_2 t_{x,t}$$

8.1.6 SOLAR RADIATION

Solar radiation is also modeled as a Gaussian with an identity link function:

$$\mathrm{R}_s = \mu + \beta_{i,1} C_t + \beta_{i,2} S_t + + \gamma_1 P_{o,t} + \zeta_1 t_{n,t-1} + \zeta_2 t_{x,t-1} + \nu_1 t_{n,t} + \nu_2 t_{x,t}$$

8.2 STOCHASTIC WEATHER GENERATOR WITH COVARIATES

With covariates, the above equations for precipitation occurrence, precipitation intensity, minimum and maximum temperature, and minimum and maximum relative humidity become:

8.2.1 PRECIPITATION OCCURRENCE

$$\ln\left(\frac{p_o}{1 - p_o}\right) = \mu + \alpha J_{t-1} + \beta_{o,1} C_t + \beta_{o,2} S_t + \gamma_1 C_t J_{t-1} + \gamma_2 S_t J_{t-1} + \varsigma U_o$$

8.2.2 PRECIPITATION INTENSITY

$$\ln(p_i) = \mu + \beta_{i,1} C_t + \beta_{i,2} S_t + + \varsigma U_i$$

8.2.3 TEMPERATURE

$$t_{n,t} = \mu_{n,0} + \mu_{n,1} + \gamma_{tn} t_{n,t-1} + \zeta_{tx} t_{x,t-1} + C_t + \beta_{tn,1} C_t + \beta_{tn,2} S_t + \eta_{t,x} P_{o,t} + \varepsilon_{tn} + \varsigma U_{tn}$$

$$t_{x,t} = \mu_{x,0} + \mu_{x,1} + \gamma_{tn} t_{x,t-1} + \zeta_{tx} t_{x,t-1} + C_t + \beta_{tx,1} C_t + \beta_{tx,2} S_t + \eta_{t,x} P_{o,t} + \varepsilon_{tx} + \varsigma U_{tx}$$

8.2.4 RELATIVE HUMIDITY

$$\ln \frac{h_n}{1 - h_n} = \mu + \alpha_1 J_{n,t-1} + \alpha_2 J_{x,t-1} + \beta_{hn,1} C_t + \beta_{hn,2} S_t + \gamma_1 P_{0,t} + \zeta_1 t_{n,t-1} + \zeta_2 t_{x,t-1} + \eta_1 t_{n,t} + \eta_2 t_{x,t} + \varsigma U_{hn}$$

$$\ln \frac{h_x}{1 - h_x} = \mu + \alpha_1 J_{n,t-1} + \alpha_2 J_{x,t-1} + \beta_{hn,1} C_t + \beta_{hn,2} S_t + \gamma_1 P_{0,t} + \zeta_1 t_{n,t-1} + \zeta_2 t_{x,t-1} + \eta_1 t_{n,t} + \eta_2 t_{x,t} + \nu h_n + \varsigma U_{hx}$$

For the equations shown in A-1.2.1 to A-1.2.4, U is the set of covariates used with each variable, and ς is the associated set of model parameters.