

Hydroclimatic prediction and projection to guide decision-making for
agricultural planning, economic development, and reservoir filling in
Ethiopia

By
Ying Zhang

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Civil and Environmental Engineering)

at the
UNIVERSITY OF WISCONSIN-MADISON
2017

Date of final oral examination: 12/21/2016

The dissertation is approved by the following members of the Final Oral Committee:

Paul J. Block (Advisor), Assistant Professor, Civil and Environmental Engineering

Steven P. Loheide, Associate Professor, Civil and Environmental Engineering

Daniel Wright, Assistant Professor, Civil and Environmental Engineering

Daniel Phaneuf, Professor, Agricultural and Applied Economics

Semu Moges, Associate Professor, Civil and Environmental Engineering

Abstract

Water is vital to life. Irregular distribution of fresh water in both time and space can cause excess and scarcity, floods and droughts, food and energy crises, water conflicts, and many other physical and social challenges. To efficiently utilize and manage water resources requires a thorough understanding of climate, hydrology, social and economic aspects in the spatial-temporal distribution of water resources, potential optimal responsive strategies, and associated social welfare. In this dissertation, the ability of hydroclimatic *prediction* to skillfully guide decision making for agricultural planning (Chapter 1, 2, and 3) and the application of hydroclimatic *projection* to inform reservoir filling (Chapter 4 and 5) are explored. Specifically, Chapter 1 addresses various clustering methods for regionalization given high spatial-temporal variation in seasonal precipitation and proposes new clustering criteria. The resulting homogenous regions are applied in Chapter 2 to produce (spatially) high-resolution season-ahead precipitation predictions. Chapter 3 focuses on understanding and quantifying the potential value of prediction in agricultural planning using an agro-economic model. Based on projections of climate variability and climate change, Chapter 4 highlights the importance of considering reservoir filling policies for the Grand Ethiopian Renaissance Dam (GERD) and implications on downstream riparian countries. Chapter 5 presents wavelet analysis to include latent hydroclimatic signals in projections and explores additional filling policies to investigate their effects on GERD hydropower production for the upstream country in addition to downstream releases.

Acknowledgements

This research was partially funded by NASA project NNX14AD30G and NSF PIRE Project 1545874. I acknowledge the National Meteorological Agency of Ethiopia for sharing data.

A lot of scientists, engineers, and economists have helped or inspired me for completing this work. Especially I submit many thanks to my advisor, Paul Block, who has provided valuable opinions and supports for my research and has also inspired me to become a better researcher and teacher. It has been a great pleasure to work with him. I would also like to thank all my committee members for their devoted time and helpful advice, Semu Moges from Addis Ababa University for generously sharing data, Ben Zaitchik and Hamada Badr from Johns Hopkins University for sharing their opinions on cluster analysis, Liangzhi You and Xinshen Diao from the International Food Policy Research Institute (IFPRI) for their help with the agro-economic model study, Michael Hammond from the University of Exeter, UK, Solomon Erkyihun from the University of Wisconsin-Madison, and Andrew King from Drexel University for their contributions on the Grand Ethiopian Renaissance Dam (GERD) work, and Donghoon Lee for his help in GIS and being a great officemate.

Finally, I thank my family and friends for their love and encouragement.

Contents

Abstract	i
Acknowledgements	ii
List of Tables	vi
List of Figures	viii
Introduction	1
Chapter 1 Optimal Cluster Analysis for Objective Regionalization of Seasonal Precipitation in Regions of High Spatial-Temporal Variability: Application to Western Ethiopia.....	6
1.1 Introduction	6
1.2 Gridded Precipitation Datasets	10
1.3 Hierarchical and Non-hierarchical Cluster Analysis	10
1.3.1 Description	10
1.3.2 Selection of k	13
1.3.3 Sources of Uncertainty	15
1.3.4 Clustering Criteria	16
1.3.5 Evaluation of Techniques on Western Ethiopian Precipitation	17
1.4 Selection of k for Non-hierarchical Cluster Analysis	23
1.4.1 Elbow Method and “Difference in minWSS”	23
1.4.2 Gap Statistic and “Difference in Difference”	26
1.4.3 Visualization of Cluster Maps	29
1.4.4 Sensitivity Analysis using Reduced Time series	31
1.5 Identification of Local and Large-scale Cluster-level Precipitation Drivers	36
1.6 Predictions	42
1.7 Conclusions and Discussion	45
Chapter 2 Does Objective Cluster Analysis Serve as a Useful Precursor to Seasonal Precipitation Prediction at Local Scale? Application to Western Ethiopia	47
2.1 Primer on Prediction Models and Cluster Analysis.....	47
2.2 Modeling High-resolution Local Seasonal Prediction.....	50
2.2.1 Cluster Analysis.....	51
2.2.2 Statistical Modeling Approach	51
2.2.3 Dynamical Modeling Approach	55
2.2.4 Performance Metrics	57

2.3 Results	58
2.3.1 Statistical Model Predictions	58
2.3.2 Dynamical Model Predictions	63
2.4 Conclusions and Discussion	65
Chapter 3 Integrating Predictive Information into an Agro-economic Model to Guide Agricultural Planning.....	68
3.1 Introduction	68
3.2 Methods	69
3.2.1 Ethiopia's Economy-wide Multi-market Model (EMM)	69
3.2.2 Climate Yield Factor (CYF) and Incorporating Climate Variability into the EMM	71
3.2.3 Integrating Seasonal Prediction to Guide Agricultural Planning – Altering Area Planted for Two Main Crops in Ethiopia	73
3.2.4 Overall Methodology Structure	74
3.3 Results	75
3.3.1 CYF and Crop Land Devoted to Maize and Teff	75
3.3.2 Incorporating Climate Variability – Baseline Simulation	77
3.3.3 Reallocating Teff and Maize Crop Land Simulations	80
3.4 Conclusions and Discussions	86
Chapter 4 Ethiopia's Grand Renaissance Dam: Implications for Downstream Riparian Countries	89
4.1 Introduction and Background	89
4.2 Nile Basin Modeling Framework	95
4.2.1 Nile Hydrology and GERD Reservoir Models	95
4.2.2 Model Data Requirements	100
4.2.3 GERD Filling Policies	100
4.2.4 Time series Development	102
4.3 Modeling Results for Streamflow at the Gezira Scheme and Laker Nasser.....	104
4.4 Summary and Discussion	112
Chapter 5 Filling the GERD: Evaluating Hydroclimatic Variability and Impoundment Strategies for Blue Nile Riparian Countries.....	114
5.1 Introduction	114
5.2 What Does the Future Hold? Projecting Hydroclimatic Conditions	115
5.3 How Will the Reservoir be Filled? Possible Filling Strategies	122
5.4 From Climate to Storage to Allocation: Model Structure	124

5.5 Results	125
5.5.1 Time to Fill the Reservoir.....	125
5.5.2 Downstream Releases during the Filling Stage.....	128
5.5.3 Hydroelectric Generation	133
5.6 Discussion	136
Chapter 6 Summary and Recommendations	138
Bibliography	140

List of Tables

Table 1-1: WSS computed for clustering results in Figure 1-5	20
Table 1-2: Inter-correlation and Intra-correlation	22
Table 1-3: intra- and inter-correlation table for k-means clustering results given $k = 8$	35
Table 1-4: Climate Variables (C.V.) in May over different regions for each cluster (C1 ~ C8) and region as a whole (non-cluster) used as predictors, with corresponding correlation between the climate variable averaged over the region and the cluster-level JJAS seasonal total precipitation time-series shown (only cells with correlation values shown are used as pre-predictors).	44
Table 1-5: Pearson correlation coefficients (Corr.) and median RPSS values corresponding to the cross-validated (CV) modeled results compared with observations in Figure 1-17	44
Table 2-1: Equations of linear regression panel models under four scenarios	55
Table 2-2: Correlation coefficients (Corr.) and RPSS for predictions (drop-one-year cross-validated) at cluster level compared to observations under C-I and NC-I scenario (extended from Table 1-5).	60
Table 2-3: Grid-level Pearson correlation and RPSS statistics	63
Table 2-4: grid-level AIC, BIC, and GCV value statistics	63
Table 3-1: Pearson correlation coefficients between JJAS total precipitation and different economic indicators over years, respectively.	78
Table 4-1: Median of 2017-2032 percent change in annual average streamflow at Gezira Irrigation Scheme for each filling policy and precipitation trend. All values are relative to the no dam policy and no changes to future precipitation.	106
Table 4-2: Median of 2017-2032 percent change in annual average streamflow at Lake Nasser for each filling policy and precipitation trend. All values are relative to the no dam policy and no changes to future precipitation.	106
Table 4-3: Median of 2017-2022 percent change in annual average streamflow at Gezira Irrigation Scheme for each filling policy and precipitation trend. All values are relative to the no dam policy and no changes to future precipitation.	107
Table 4-4: Median of 2017-2022 percent change in annual average streamflow at Lake Nasser for each filling policy and precipitation trend. All values are relative to the no dam policy and no changes to future precipitation.	107

Table 5-1: Time to fill the reservoir for different filling strategies.	127
Table 5-2: lower 95 th percentile and median values of annual GERD releases (1, 2 and 3-year shortfalls in km ³) for each filling strategy; 24 months and 36 months are per year averages.	130
Table 5-3: Annual hydroelectric generation (GW-hrs) during the filling stage (unique to each strategy) and full projection period (2018 - 2034), assuming full hydroelectric generation after the reservoir is filled (15 000 GW-hrs annually). Calculations are based on the mean values of 100 simulations and numbers are rounded to hundreds.	135
Table 5-4: Discounted present value (as of December 2014) of annual benefit (million USD) from hydroelectricity generated during the filling stage (unique to each strategy) and full projection period (2018 – 2034), assuming electric price of \$0.07/kW-hrs, 3% annual discounting rate and full hydroelectric generation after the reservoir is filled (15 000 GW-hrs annually). Calculations are based on the mean values of 100 simulations.	135

List of Figures

Figure 0-1: Dissertation framework with research questions addressed in each chapter.	5
Figure 1-1: Study region (framed) of western Ethiopia and sample sites with June-September seasonal total precipitation (mm) time series from 1983 to 2011. Circles indicate the precipitation in 1997 which is a strong El Niño year.	8
Figure 1-2: Synthetic hierarchical clustering dendrogram conceptualization initialized with ten data units. At the cutoff point shown, three clusters are identified.	11
Figure 1-3: Synthetic k-means clustering conceptualization based on a 2-dimentional dataset (X, Y); the diamond and the cross represent the data point and the centroid, respectively: (a) scatter plot of the data; (b) randomly assigned centroids to initiate the algorithm; data points assigned to the closest centroid (labeled with the same index as its assigned centroid); (c) re-calculated centroids and re-assigned data points; (d) re-calculated centroids and same re-assigned results (convergence).	12
Figure 1-4: WSS, given different number of clusters k based on k-means clustering results on JJAS seasonal total precipitation over the complete study region. <i>Note that, for each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually in crosses.</i>	14
Figure 1-5: Comparative Results of hierarchical clustering specifically using (a) Regional Linkage, (b) Average Linkage and (c) Ward's method as distance criteria and non-hierarchical (d) k-means clustering over (1) the region of 5.15N – 14.95N, 33.05E – 39.95E (top row), and (2) with additional region of 4.15N – 5.15N, 34.95E – 38.75E (bottom row). All with the number of clusters $k = 5$	19
Figure 1-6: K-means clustering map under smoothing factor of 5 (top left), 10 (top right), 15 (bottom left) and 25 (bottom right)	20
Figure 1-7: Minimum WSS, AIC, BIC and GCV given different number of clusters k based on k-means clustering results on JJAS seasonal total precipitation over the complete study region. ...	25
Figure 1-8: The minWSS and difference in minWSS from $k-1$ to k	26
Figure 1-9: Gap statistic and “difference in difference” results.	28
Figure 1-10: K-means clustering maps given different number of clusters k ranging from 2 to 10	30

Figure 1-11: Mean time series of JJAS seasonal total precipitation over the complete study region.	32
Figure 1-12: clustering maps with full time series (top left), drop 5 driest years (bottom left), drop first 5 years (top right), and drop last 5 years (bottom right) for $k = 7, 8$ and 9	33
Figure 1-13: k-means clustering maps given $k = 8$ with country boundary and river profile.	34
Figure 1-14: Standardized mean time series (centroid) of JJAS seasonal total precipitation within each clustered region from 1983 to 2011 (29 years); clustering results are based on $k = 8$	35
Figure 1-15: Global correlation map of (a) SST and (b) SLP correlated with within-cluster mean time series of JJAS seasonal total precipitation.	40
Figure 1-16: Correlation map of (a) surface air temperature, (b) outgoing longwave radiation correlated with within cluster mean time series of JJAS seasonal total precipitation for (a) Cluster 5 and 7, (b) Cluster 2 and 6, centered at Africa, with dashed square indicating the study region and solid line circling the corresponding cluster region.	41
Figure 1-17: Modeled results using the first two PCs of selected climate variables as predictors, with drop-one-year cross-validation.	43
Figure 2-1: Flow chat of data processing for predictors into the statistical model. Numbers framed by dash lines correspond to the procedures listed in the context. <i>Note: pre. – precipitation, t-s – time series, avg. – average.</i>	53
Figure 2-2: (a) bias correction of NMME predictions using probability mapping; (b) precipitation time series from NMME (colored lines) before and after correction, compared to observations (black line). Examples are shown for randomly selected six grid-cells.	57
Figure 2-3: cluster-level predictions and observations under C-I and NC-I scenario, with drop-one-year cross-validation. The 95% envelope shows the 95% confidence interval constructed using model errors (extended from Figure 1-17).	59
Figure 2-4: Pearson correlations between grid-level observations and predictions under four scenarios, with the clustering boundary delineated roughly in black.	61
Figure 2-5: grid-level RPSS (%) under four scenarios using climate variables as predictors, with the clustering boundary delineated roughly in black.	62
Figure 2-6: Boxplots of grid-level RPSS (%) for 10 dynamical models from NMME (a) before and (b) after bias correction, labeled with the same number as listed in the context. <i>Note: For each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, and the whiskers extend to the most extreme data points not considered outliers (outliers not shown).</i>	64

Figure 2-7: Pearson correlations between grid-level observations and ensemble mean of bias-corrected predictions for 10 dynamical models from NMME, labeled with the same number as listed in the context. <i>Note that the scale ranges from -1 to 1.</i>	65
Figure 3-1: Flow chat of methodology. Solid line arrows indicate the process of baseline simulation and dashed line arrows indicate the add-on prediction-guided simulations.	75
Figure 3-2: Spatial average zonal CYF values over all 56 zones for six staple crops, with barley assumed to share the same CYF values as sorghum.	76
Figure 3-3: Crop area allocation in Ethiopia. “PulsesOther” represents pulses except beans and peas. “OilcropsOther” represents oil crops except sesame, groundnut, and rape seed.....	77
Figure 3-4: National real GDP and spatial average JJAS total precipitation over simulated years.	78
Figure 3-5: Total calorie per capita per day (Cal) averaged over food deficit zones and food surplus zones, respectively.	79
Figure 3-6: Percent of poor population over total population in each category (rural, urban, and total) over the simulated years.....	80
Figure 3-7: Real GDP change from baseline for all years with above-normal, near-normal, and below-normal conditions under (a) teff-to-maize and (b) maize-to-teff scenarios, averaged over outputs with different area reallocating percentages (5%, 10%, ..., 95%).....	81
Figure 3-8: Seasonal prediction of JJAS precipitation from 1983 to 2011.	82
Figure 3-9: Real GDP change from baseline under “teff-to-maize” scenario (a) averaged over outputs with different area reallocating percentages (5%, 10%, ..., 95%), and (b) for each reallocating percentage respectively. <i>Note that years with above-normal condition predicted are marked with crosses (prediction-guided outputs).</i>	84
Figure 3-10: Calorie per capita per day (Cal) change from baseline averaged over predicted above-normal years for different teff-to-maize area reallocation percentages (prediction-guided outputs).....	85
Figure 3-11: Import quantity of teff for predicted above-normal years across different teff-to-maize area reallocation percentages (import of teff is triggered starting at 55%).	85
Figure 3-12: Poor population change (rural and urban) and poverty rate change from baseline averaged over action-adopted years.	86
Figure 4-1: The Nile basin with the forthcoming GERD. Base map courtesy of World Wildlife Fund (WWF, 2007)	90

Figure 4-2: Contributing rivers to the main Nile discharge at Lake Nasser using monthly discharge data from 1961-1985 (for Atbara River and Blue Nile) and 1961-1995 (for White Nile) (Adapted from Hurst 1952).	91
Figure 4-3: Schematic of the model framework.....	96
Figure 4-4: Nile WatBal hydrology model recalibration on monthly streamflow at (a) Roseires (b) Khartoum and (c) Dongola; (d) Monthly streamflow averaged over 1961-1995 and monthly cumulative error in volume of modeled and observed streamflow.	98
Figure 4-5: Flowchart of methodology.....	99
Figure 4-6: GCM simulations of historical and projected precipitation changes as a percent of the 1950-1990 average annual precipitation. The multi-model mean projection is represented by the bold line. GCMs are from the International Panel on Climate Change Fourth Assessment Report (IPCC 4AR) (Pachauri and Reisinger, 2007) based on the A1B scenario from 2000-2100 (Adapted from Giannini et al., 2008).	103
Figure 4-7: Percent reduction in annual streamflow entering Lake Nasser considering no future changes in precipitation. Three filling policies shown. <i>Note: For each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually in crosses.</i>	105
Figure 4-8: Box plots of 2017-2032 percent change in annual average streamflow at Lake Nasser for each filling policy and precipitation trend constructed from the 100 time series. All values are relative to the median no dam policy and no changes to future precipitation. <i>Note: For each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually in crosses.</i>	109
Figure 4-9: Box plots of 2017-2032 percent change in annual average streamflow at (a) Gezira Scheme and (b) Lake Nasser for each filling policy constructed from the 100 time series and weighted precipitation changes. All values are relative to the no dam policy and no changes to future precipitation. <i>Note: For each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually in crosses.</i>	111
Figure 4-10: Box plots of 2060 percent change in annual average streamflow at Lake Nasser for the 25% filling policy and precipitation trend constructed from the 100 time series. <i>Note: For each box plot, the short line inside the box is the median, the box edges represent the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually in crosses.</i>	111
Figure 5-1: Blue Nile basin precipitation: (a) monthly average, and (b) annual total based on CRU TS v3.23 data set (https://crudata.uea.ac.uk/cru/data/hrg/)	116

- Figure 5-2: Wavelet analysis of Blue Nile annual precipitation: (a) wavelet power spectrum, (b) global wavelet spectrum, with dashed lines showing the 90% (left) and 95% (right) significance levels and (c) standardized detrended annual precipitation (black solid line), with combined low-frequency signal reconstructed at 21-year and 35-year periodicity, scaled. 117
- Figure 5-3: Illustration of the Block-KNN projection methodology. 119
- Figure 5-4: (a) Aggregated low frequency signal in Blue Nile annual precipitation over 1901-2014 ($\alpha_1 S_{21t} + \alpha_2 S_{35t}$; see Equation 5-1); values are scaled. (b) Residual annual time series after subtracting the signal in (a) from the detrended normalized historical record. (c) Residual monthly time series after disaggregating from annual time series to monthly time series and subtracting the signal in (a) using Equation 5-2. 121
- Figure 5-5: Scaled monthly precipitation from the historical and projected periods, highlighting the embedded low frequency signal (single simulation only). 121
- Figure 5-6: Volume of water in the reservoir over time for different filling strategies based on projections with the low frequency signal embedded. Solid and dashed lines indicate the median and lower 95th percentile, respectively, based on the 100 simulations. 126
- Figure 5-7: Annual downstream releases from the dam for different filling strategies based on projections with the low frequency signal embedded. Solid and dashed lines indicate the mean and lower 95th percentile, respectively, based on the 100 simulations. 128
- Figure 5-8: Percentage reduction in annual downstream releases for each filling strategy, compared with historical averages; the low frequency signal is embedded. Boxplots are based on 100 simulations. *Note: For each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, and the whiskers extend to the 5th and 95th percentiles.* 130
- Figure 5-9: Number of months with no downstream releases for the three absolute filling strategies based on projections with the low frequency signal embedded, accumulated from the 100 simulations. *Note: the y-axis scale differences for each subplot.* 131
- Figure 5-10: Number of months with no flow released downstream of the dam for three absolute filling strategies, accumulated for all months in each simulation (ordered). Solid and dash-dot lines indicate the results based on projections with and without the low frequency signal embedded, respectively. 133
- Figure 5-11: Annual hydroelectric generation based on projections with the low frequency signal embedded. Solid and dashed lines indicate the mean and lower 95th percentile of 100 simulations, respectively. 134

Introduction

Prediction and projection of hydroclimatic variables, such as precipitation and streamflow, involve both climatic and hydrological processes. Hydroclimatic *prediction* refers to a relatively short time scale, ranging from hourly to annually, while *projection* addresses long-term changes, often simulating general trends for decades into the future.

Precipitation prediction at seasonal scales aims to provide actionable information to guide planning and management of various sectoral activities. Advanced skill in seasonal precipitation prediction may offer decision makers opportunities to realize benefits given expected advantageous conditions, while reducing unnecessary costs for anticipated disadvantageous situations. It is potentially useful for developing countries such as Ethiopia, where reliance on rain-fed agriculture and minimal water resources infrastructure make it vulnerable to precipitation fluctuations. However, producing skillful seasonal predictions in Ethiopia is challenging, particularly at local scales, as precipitation is highly varied temporally and spatially (e.g. Gissila et al., 2004; Block and Rajagopalan, 2007).

Operational precipitation predictions in Ethiopia have been issued by its National Meteorological Agency (NMA) since 1987 using an analog methodology (i.e. locating a similar climate scenario in the past – an analog – to predict future conditions); however, this approach has produced only marginally skillful outcomes (Korecha and Sorteberg, 2013). For NMA's prediction, the country is divided into eight homogeneous regions for which NMA produces independent predictions. Similarly, others have also addressed

seasonal prediction in Ethiopia contingent on both temporal and spatial precipitation patterns (e.g. Gissila et al., 2004; Block and Rajagopalan, 2007; Korecha and Barnston, 2007; Diro et al., 2009; Segele et al., 2015). However, all of these studies focus on predicting regional average precipitation based on subjective clustering methods applying a limited number of stations or coarsely gridded data; no local predictions at a finer spatial scale have been explored (see detailed literature review in Chapter 1 and 2).

Hence, Chapter 1 and 2 address the two questions below, respectively – **“How can homogeneous regions for seasonal precipitation prediction be rigorously defined?”** and **“Given these homogeneous regions, can fine-scale seasonal prediction be improved?”**

For agricultural planning, particularly rainfed crops, yields are often highly sensitive to climate conditions; thus, given a climate prediction, potential benefits to farmers may be possible through increased food production, improved investment efficiency, and reduced risks. Evidence from farmer surveys and interviews regarding their perspectives and actual use of predictions and associated outcomes suggest a demand for prediction and potential benefit of applying climate predictions (e.g. Patt et al., 2005; Roncoli, 2006; Roncoli et al., 2009); however, a notable gap still exists between scientific researchers who produce prediction and the end users of the predictive information. Factors related to unfit prediction products, communication failure, lack of governmental or institutional support, limited access to information, minimal capacity to respond, and data scarcity have constrained the widespread use and benefit of seasonal prediction, particularly for

smallholder farmers in less-developed countries (Broad and Agrawala, 2000; Hansen, 2002; Hansen et al., 2011).

An interdisciplinary approach connecting climate prediction to agricultural planning could help to communicate information in a manner more relevant to decision makers. An *ex-ante* evaluation of climate predictions can provide quantitative measurement of anticipated benefits given a range of possible response strategies – a potentially more concrete image than simply interpreting climate predictions. Correspondingly, the gap between prediction and its application may be alleviated by the *ex-ante* evaluation of seasonal prediction in two general ways (Thornton, 2006; Meza et al., 2008). First, it can provide motivation to mobilize funds and influence the agendas of institutional partners when facing competing priorities. Second, it can inform policy and decision makers by providing insights on investments in specific response strategies and their associated net benefits.

Chapter 3 addresses the challenge of alleviating the gap between prediction and its application by asking **“What is the economic value of seasonal climate prediction for Ethiopian agricultural planning?”**

Hydroclimatic information is not limited to the Ethiopian agricultural sector of course, but also has important implications for large-scale water resources management in Ethiopia and other riparian countries who share the Blue Nile River. The Grand Ethiopian Renaissance Dam (GERD), currently under construction in western Ethiopia on the Blue Nile River, is soon to be Africa’s largest dam. It is considered an important part of the

strategic hydropower development plan in Ethiopia – often referred to as the Africa’s water tower – to serve domestic and foreign electricity demands and address regional economic water scarcity (WWAP, 2012). Developing water infrastructure in Ethiopia may help to reduce poverty throughout the region. While there have been numerous recent studies debating the long-term economic, sectoral, resource, and social implications of the GERD (e.g. Mulat and Moges, 2014; Arjoon et al., 2014; Bastawesy, 2014; Abdelhady et al., 2015; Kahsay et al., 2015), less attention has been devoted to potential impacts on riparian countries during the reservoir filling stage. Given the massive 74 km³ reservoir volume and large interannual variations in Blue Nile streamflow, the manner in which the reservoir is filled is critical. Ethiopia has incentive to fill the reservoir rapidly to begin generating hydropower, while downstream countries prefer a modest filling rate to minimize the possible impacts on dependent lives, livelihoods, and ecosystems.

No policy dictating the reservoir filling policy has been publicly agreed upon by all riparian countries, raising the question **“What are the implications of various GERD reservoir filling policies on downstream flows in Sudan and Egypt and upstream hydropower generation in Ethiopia?”** Through exploring a range of reservoir filling policies associated with their implications to both downstream and upstream riparian countries based on hydroclimatic projections, the two chapters aim to provide regional decision makers with a set of plausible, justifiable, and comparable outcomes in order to foster agreeable mutual policies.

A roadmap of research questions and associated chapters is shown below:

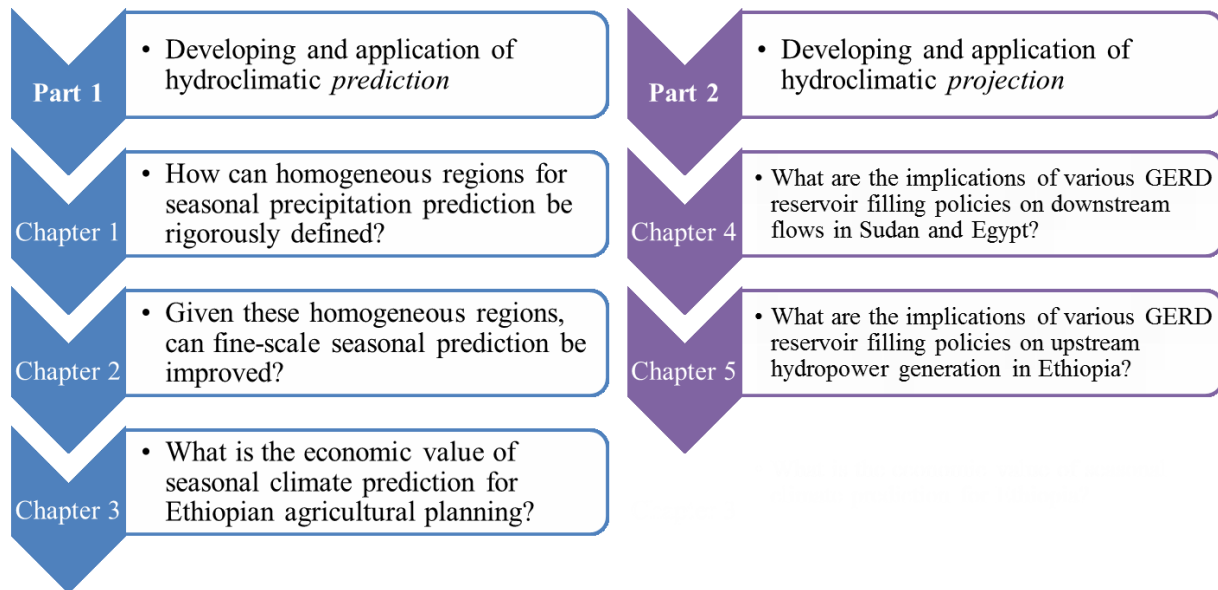


Figure 0-1: Dissertation framework with research questions addressed in each chapter.

Chapter 1 Optimal Cluster Analysis for Objective Regionalization of Seasonal Precipitation in Regions of High Spatial-Temporal Variability: Application to Western Ethiopia

1.1 Introduction

Defining homogeneous precipitation regions for hydrologic modeling, ecological and climate classification, prediction or other analysis is non-trivial given variation in both temporal and spatial patterns. Multiple methods exist to delineate boundaries and define the optimal number of clusters (e.g. Bisetegnè et al., 1986; Gong and Richman, 1995; Jain et al., 1999; Gissila et al., 2004). Ideally an objective method is selected to foster reproducibility; however even so, traditional approaches typically include aspects of subjective delineation. Here we evaluate various regionalization methods for objective delineation and define a number of approaches for optimally selecting an appropriate number of clusters. These techniques are applied to seasonal precipitation in western Ethiopia for illustration; however, transferability to other variables and regions is possible.

Precipitation in western Ethiopia is tied to many important sectors, defining lives, livelihoods, and major parts of the domestic economy. It is the source of the Blue Nile River and others, offering substantial hydropower potential, second only to the Democratic Republic of Congo (Bartle, 2002). Also, given that only 1% of the cultivated land is irrigated (Korecha and Sorteberg, 2013), rain-fed yields are chiefly subject to precipitation quantity and timing, effectively dictating Ethiopia's agriculture economy. Precipitation extremes – both droughts and floods – are also not uncommon across the

country, exacerbating Ethiopia's vulnerability. Improving our understanding of Ethiopia's interannual variability in precipitation could benefit the country by developing a reliable seasonal prediction system to improve strategic planning of agriculture production, reservoir operations and water allocation, and preparation for potential natural disasters. However, with high temporal variability, this is a challenging task.

In addition to large interannual variability (Figure 1-1), highly variable spatial patterns of precipitation also add complexity to attribution and prediction. Numerous studies to date point to the migration of the Inter-tropical Convergence Zone (ITCZ), multiple regional hydroclimatic system interactions, and topographic influences as the leading explanatory mechanisms in describing precipitation variability in Ethiopia (e.g. Griffiths, 1972; Gamachu, 1977; NMSA, 1996; Conway, 2000; Seleshi and Zanke, 2004). Tele-connected large-scale climate variables are also shown to be influential, particularly the El Niño Southern Oscillation (ENSO) (e.g. NMSA, 1996; Bekele, 1997; Camberlin, 1997; Wolde-Georgis, 1997; Gissila et al., 2004; Segele and Lamb, 2005; Block and Rajagopalan, 2007; Korecha and Barnston, 2007; Diro et al., 2011a; Elagib and Elhag, 2011). More recently, effects of the Indian Ocean are being cited (e.g. Shanko and Camberlin, 1998; Goddard and Graham, 1999; Latif et al., 1999; Black et al., 2003), as are the Azores, St. Helena, and Mascarene high pressure systems (Kassahun, 1987; Tadesse, 1994; NMSA, 1996; Segele and Lamb, 2005). These numerous and diverse drivers of variability, and their interactions, lead to a complex spatial and temporal precipitation regime. Some efforts at regionalization – specifically defining boundaries of homogeneous precipitation regions – have been undertaken, but traditionally rely on subjective delineation.

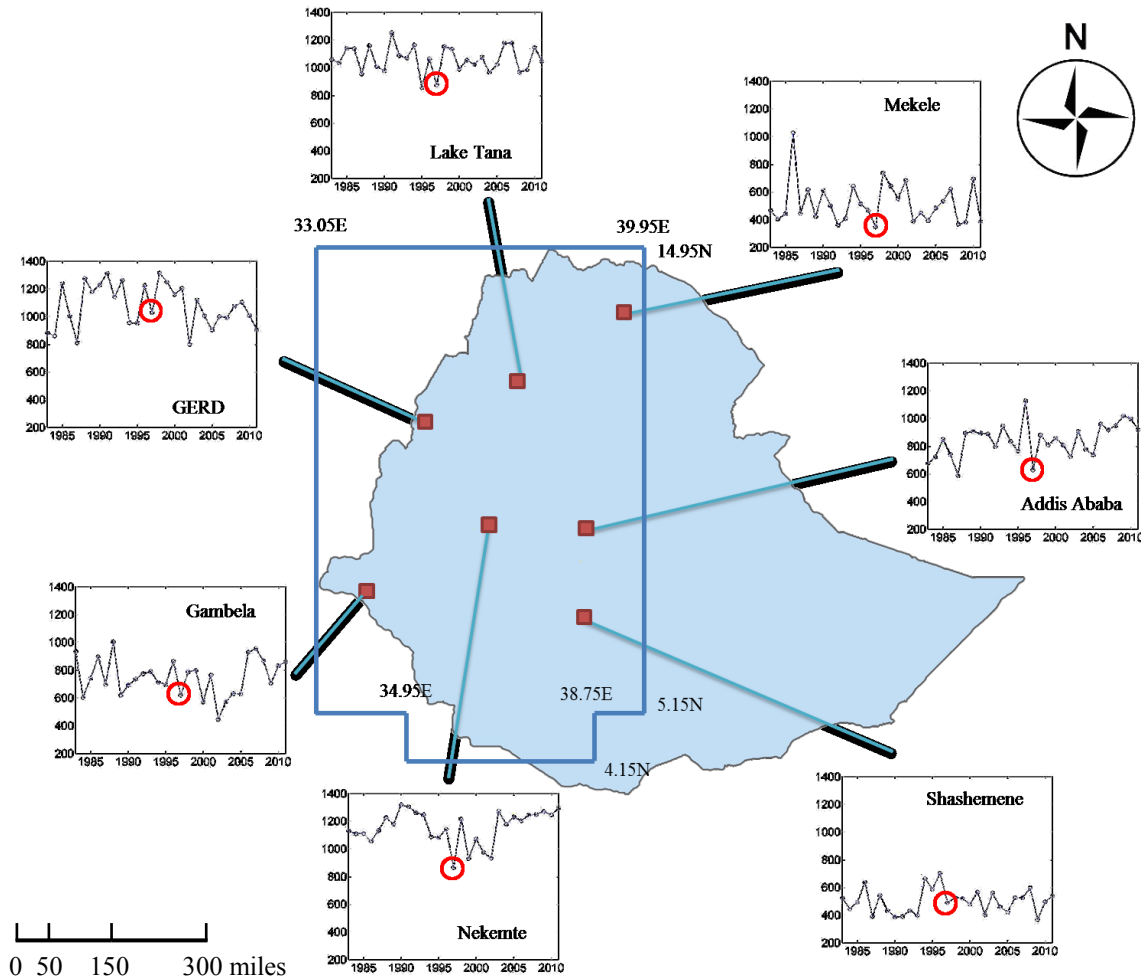


Figure 1-1: Study region (framed) of western Ethiopia and sample sites with June-September seasonal total precipitation (mm) time series from 1983 to 2011. Circles indicate the precipitation in 1997 which is a strong El Niño year.

Currently, the National Meteorological Agency (NMA) of Ethiopia divides the country into eight homogenous regions subjectively according to major atmospheric and oceanic circulation mechanisms and typical rain-producing systems affecting respective regions (Korecha and Sorteberg, 2013). A number of research studies have also proposed regionalization methods and subsequent clusters. Gissila et al. (2004) group Ethiopia into four clusters by comparing the seasonal cycle subjectively and analyzing the coherence of

interannual variability from 24 stations. Diro et al. (2009) follow the same methodology but divide Ethiopia into five regions based on data from 33 stations; in addition, they adjust regional boundaries according to the interannual variability for spring (February-May) and summer (June-September) season respectively, considering different homogeneous regions affected by diverse large-scale forcings in different season. In other studies, principal component analysis (PCA) is applied to identify homogeneous rainfall zones. Bisetegn et al. (1986) create five regional groups based on only 21 Ethiopian rainfall stations using PCA, retaining four eigenvectors explaining 75% of the variance. Eklundh and Pilesjö (1990) also perform PCA on rainfall data from 63 stations and divide the country into seven regions. These methods were all applied on non-uniform station data and require *subjective* grouping of stations, interpretation, and hand-delineation of boundaries. Not only are these outputs time-consuming to develop, the underlying methods produce immeasurable subjective errors.

This motivates analysis to objectively and automatically define homogeneous precipitation zones, preferably with a uniform dataset; for this chapter, we propose a k-means clustering statistical method applied to a gridded rainfall dataset. Western Ethiopia is clustered into homogeneous regions based on the *Kiremt* season total precipitation spanning June through September (JJAS). This season produces approximately 70% of the upper Blue Nile basin annual precipitation (Conway, 2000) and coincides with the major agricultural activities (Degefu, 1987). A cluster-based season-ahead prediction model is also presented for demonstration of utility; however, seasonal prediction is not the only application of regionalization. The clustering results can also be used for regional planning and management, hazard evaluation and so forth. Hence, the objective

of this chapter is not centered on prediction techniques, but rather regionalization through cluster analysis that may subsequently lead to improved seasonal precipitation prediction.

1.2 Gridded Precipitation Datasets

A 0.1×0.1 degree gridded monthly precipitation dataset from NMA is utilized in this research. The data is a merged product of satellite estimates and station measurements with spatial coverage over western Ethiopia from 1983 to 2011 (Figure 1; Dinku et al., 2014). This product has been shown to reproduce station data over areas with both densely and sparsely distributed station networks. Data is aggregated to JJAS seasonal total precipitation over the 29 years for each of the 7320 grid-cells.

1.3 Hierarchical and Non-hierarchical Cluster Analysis

1.3.1 Description

Two main types of clustering algorithms exist for analysis of gridded data, namely hierarchical and non-hierarchical (Jain et al., 1999); the objective is to assign each grid-cell to a cluster based on the inter time series correlation given the spatial-temporal dataset, analogous to using the Euclidean distance for a 2-dimensional matrix. A higher correlation is equivalent to a smaller Euclidean distance, and dictates which grid-cells are likely to be grouped together. Hierarchical clustering produces a cluster dendrogram (Figure 1-2), where based on differently structured criteria, the two most similar grid-cells would be grouped into one branch for the first step and subsequently viewed as one

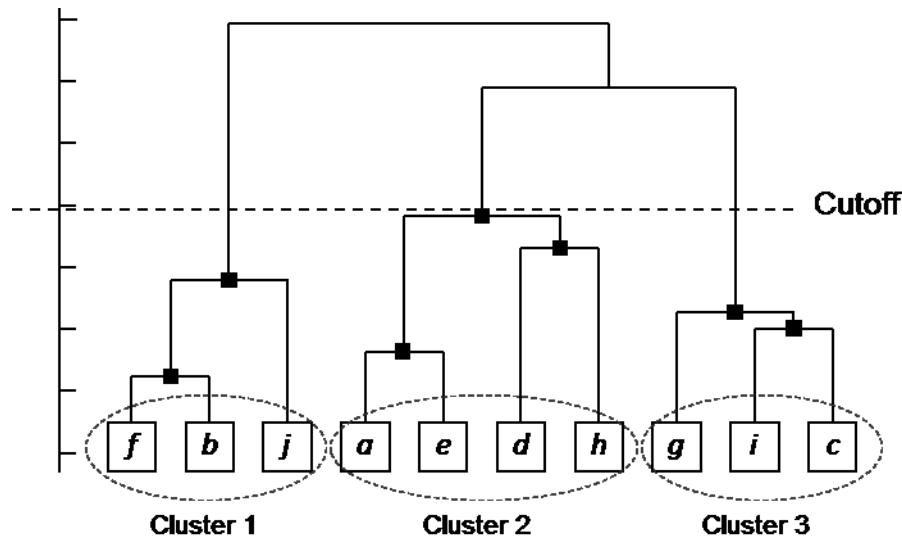


Figure 1-2: Synthetic hierarchical clustering dendrogram conceptualization initialized with ten data units. At the cutoff point shown, three clusters are identified.

new unit with an averaged time series entering the next step. Two units are combined at each step. Once a unit is assigned to a branch, it cannot be detached and the algorithm continues until the last two units are combined. Where this hierarchical “tree” is cut determines the final number of clusters. Non-hierarchical clustering, often referred to as k-means clustering, is more flexible. In contrast to hierarchical clustering, non-hierarchical clustering algorithms allow a grid-cell to be reassigned to reach an optimum result. Non-hierarchical (K-means) clustering algorithms typically follow these steps (Figure 1-3):

- (1) pre-select the number of clusters (k)
- (2) randomly pick centroids or time series of each cluster
- (3) assign grid-cells to the most similar centroid
- (4) recalculate centroids by averaging all time series assigned to that centroid
- (5) iterate steps (3) and (4) until convergence.

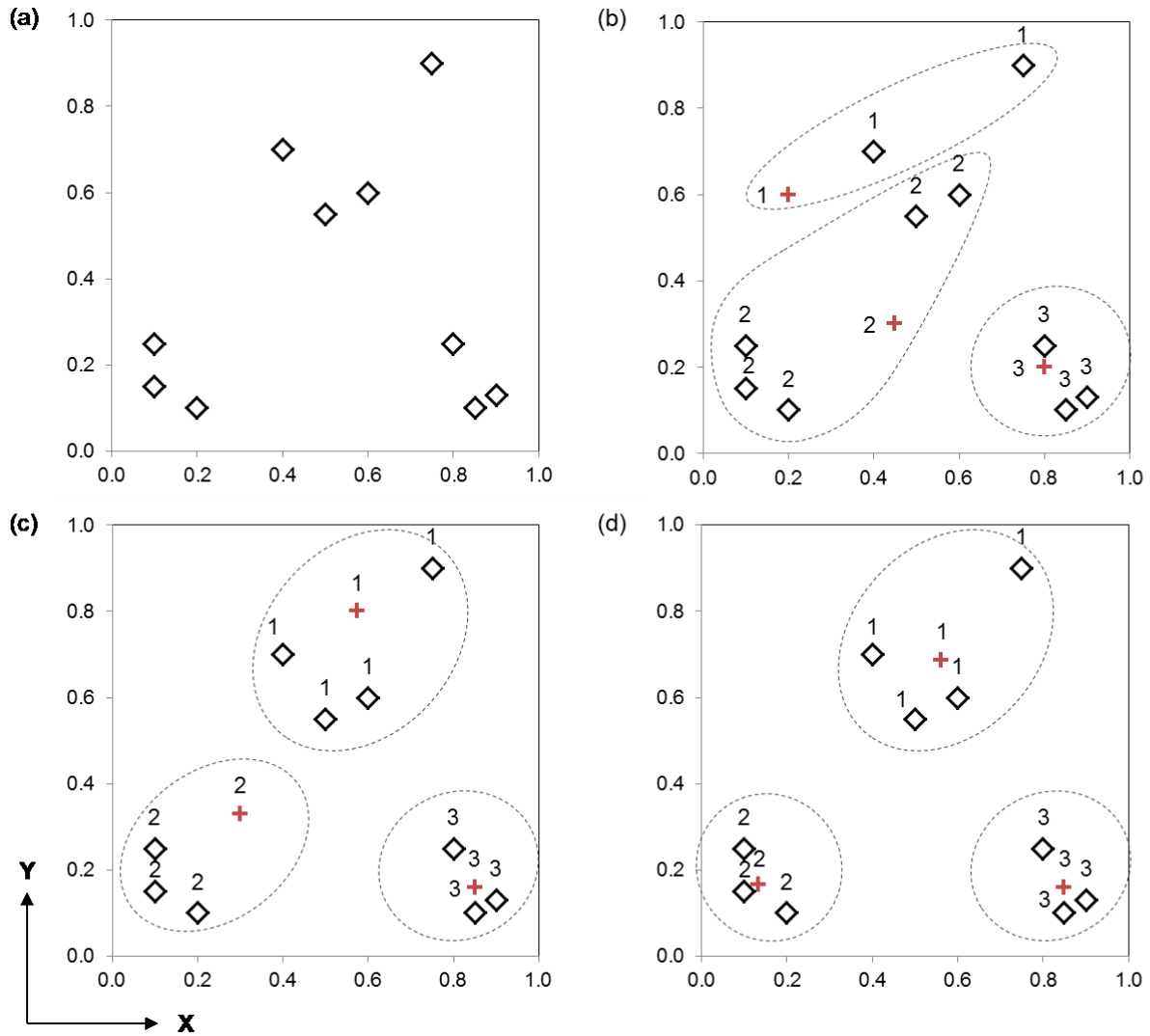


Figure 1-3: Synthetic k-means clustering conceptualization based on a 2-dimensional dataset (X, Y); the diamond and the cross represent the data point and the centroid, respectively: (a) scatter plot of the data; (b) randomly assigned centroids to initiate the algorithm; data points assigned to the closest centroid (labeled with the same index as its assigned centroid); (c) re-calculated centroids and re-assigned data points; (d) re-calculated centroids and same re-assigned results (convergence).

This produces k clusters with k cluster means or centroids and the within cluster sum of square errors (WSS) (Equation 1-1) will be minimized:

$$WSS = \sum_j \sum_{g \in j} (t_g - \bar{t}_j)^2 \dots\dots\dots \text{(Equation 1-1)}$$

where WSS is the sum of the squared errors between the time series in each grid-cell g (t_g) in cluster j ($g \in j$) and the average time series of cluster j ($\bar{t}_j$, known as the mean or centroid), and then summed over all k clusters.

1.3.2 Selection of k

Although data processing is automatic and objective for both hierarchical and non-hierarchical clustering analysis, determining k is still subjective. For hierarchical clustering, this requires a cutoff point or a desired number of clusters after the dendrogram is formed. For non-hierarchical clustering, a predetermined number of clusters is required in order to initiate the algorithm. An optimal number of clusters, however, can be estimated given a certain confidence level desired for intra-correlations (i.e. correlation between the centroid of each cluster and its members) or inter-correlations (i.e. correlation between centroids of two different clusters) (Badr et al., 2015). Intuitively, relatively high intra-correlation and low inter-correlation are desired. This estimation strategy is significantly more suitable for hierarchical clustering, as it has a fixed “tree” structure regardless of the number of clusters selected post-analysis; although non-hierarchical clustering requires a predetermined number of clusters, intra- and inter-correlations can still be examined for several trials of k .

An alternative method for selecting an optimal number of clusters is to perform a sensitivity analysis of WSS given different k (Figure 1-4), particularly for non-hierarchical clustering. By evaluating the improvement of WSS when one cluster is added, the potentially optimal k can be identified; this is known as the Elbow method

(Thorndike, 1953). This method may be problematic, however, when considering a large number of grid-cells yet a relatively small number of clusters is desired. This context is common in the field of hydroclimatology, particularly as data resolution increases yet homogeneous climatic zones remain the same size. Similar methods for determining the

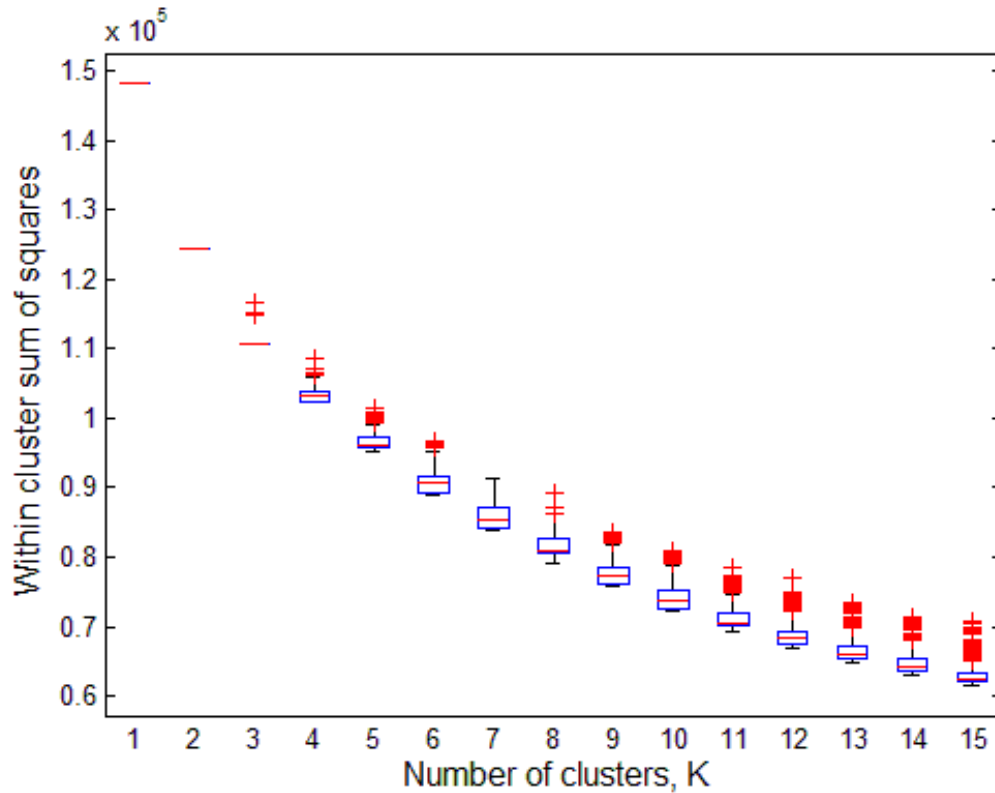


Figure 1-4: WSS, given different number of clusters k based on k-means clustering results on JJAS seasonal total precipitation over the complete study region. *Note that, for each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually in crosses.*

optimal number of clusters based on the WSS include Akaike information criterion (AIC), Bayesian information criterion (BIC), and generalized cross-validation (GCV), as well as some more sophisticated methods such as the Gap Statistic (Tibshirani et al., 2001) and the Jump method (Sugar and James, 2003), however these all suffer from the same

problem. Thus new methods or extensions of these are warranted and developed here to identify a reasonable number of clusters.

1.3.3 Sources of Uncertainty

In addition to uncertainties in methods for selecting the optimal number of clusters, additional sources of uncertainty effecting cluster outputs include noise in the raw data, the data extent, and the initial cluster state that may bias results. In hierarchical clustering, once two units are assigned to the same branch, and since reassignment is disallowed, a small bias leads to large biases in subsequent steps through a chain reaction. Comparatively, since non-hierarchical clustering is a flat divisive process, it is likely to be less effected by noisy data, if at all. If a subset of the data is initially selected and regionalized, subsequently extending the analysis spatially can be problematic particularly for hierarchical clustering. Therefore both hierarchical and non-hierarchical approaches should be subjected to a sensitivity test on spatial data extended from the initial analysis. It should also be noted that non-hierarchical clustering often suffers from non-exclusive convergence when subjected to different initial states (i.e. different cluster outcomes for different initials states). Nevertheless, by iteratively exploring different initial states, the optimal clustering result defined by the minimum WSS (minWSS) can be located.

1.3.4 Clustering Criteria

Three criteria for hierarchical clustering are selected: Ward's method, average linkage, and regional linkage for comparison with non-hierarchical (k-means) clustering. Both the Ward's method and k-means clustering minimize the WSS using the "Euclidean" distance, defined as one minus the Pearson correlation between the time series of any two units (Equation 1-2) (Ward Jr, 1963; Hartigan, 1975). Average linkage is another commonly used criterion, which links each candidate unit in the current step to two units merged in the previous hierarchical level, separately, to calculate the overall updated Pearson correlation distance weighted by the number of grid-cells in each unit (Equation 1-3). Regional linkage is an adjusted version of average linkage criterion incorporating the standard deviation (Equation 1-4) (Badr et al., 2015). The equations for calculating the distance based on different clustering criteria discussed above are listed here:

$$d_{x,y} = 1 - r_{x,y} \dots\dots\dots \text{(Equation 1-2)}$$

where $d_{x,y}$ is the equivalent Euclidean distance between x and y , based on $r_{x,y}$ which is the Pearson correlation of the two time series of units x and y ;

$$d_{x \cup y, z} = \frac{n_x d_{x,z} + n_y d_{y,z}}{n_x + n_y} \dots\dots\dots \text{(Equation 1-3)}$$

where $d_{x \cup y, z}$ is the overall updated Pearson correlation distance between candidate unit z and already merged units x, y in the previous hierarchical stage, and the number of members in units x and y are represented by n_x and n_y ; and

$$d_{x \cup y, z} = \frac{n_x \left(\frac{\sigma_x}{\sigma_{x \cup y}} \right) d_{x,z} + n_y \left(\frac{\sigma_y}{\sigma_{x \cup y}} \right) d_{y,z}}{n_x + n_y} \dots\dots\dots \text{(Equation 1-4)}$$

where $\sigma_{x \cup y}$ is the standard deviation of the mean time series of merged x and y at the final stage; all other parameters are the same as in Equation 1-3.

Since data are standardized beforehand, all of the time series in each grid-cell have a mean of 0 and variance of 1, therefore the clustering is not affected by differences in mean or variance, but instead based on the correlation among all standardized time series. Variability from one time step to the next (up or down from one specific year to the next year) plays the major role in determining the clustering results.

1.3.5 Evaluation of Techniques on Western Ethiopian Precipitation

An R-tool for hierarchical climate regionalization (Badr et al., 2015) is used to produce hierarchical clustering results based on preprocessed data initially over the main region and then extended to include the additional southernmost portion of data (Figure 1-1). This is compared with non-hierarchical k-means clustering results on the same regions. All data are preprocessed by standardizing across years for each gridded time series. An optimal number of clusters $k = 5$ is obtained at the 99% confidence level ($\alpha = 0.01$) using the regional linkage hierarchical clustering technique for both data regions. However, for non-hierarchical clustering the sensitivity analysis of k versus WSS for 1000 trials per k produces a smooth scree plot with no apparent elbow to identify the optimal number of clusters (Figure 1-4). Before further evaluation of an optimal k , comparative results of hierarchical and non-hierarchical clustering for $k = 5$ are performed to identify the preferred clustering method, if possible. Note that for each k ,

the optimal k-means clustering result is selected corresponding to the minWSS obtained from the 1000 trials with random initial states.

To evaluate the effect of extending the data analysis region, the hierarchical and non-hierarchical cluster techniques are applied initially to approximately 95% of the data (southern portion omitted) as well as the full dataset, for comparison of how cluster delineations may change (Figure 1-5). As expected, regional linkage and average linkage induce similar clustering results given their comparable structured criteria. Similarly, the Ward's method and k-means clustering results resemble each other given their equivalent objective of minimizing the WSS. K-means clustering provides the lowest WSS, followed by the Ward's method (Table 1-1). When comparing the two different data extents for each method, the differences in cluster boundaries for hierarchical clustering are quite notable, while k-means clustering basically produces the same cluster delineations. Thus k-means clustering is likely more robust than hierarchical clustering in terms of cluster boundary stability considering inevitable data noise; this ability is attributable to its flexible algorithm of assigning and reassigning grid-cells to clusters. However, this is also why it produces less smooth cluster borders compared with the hierarchical cluster methods.

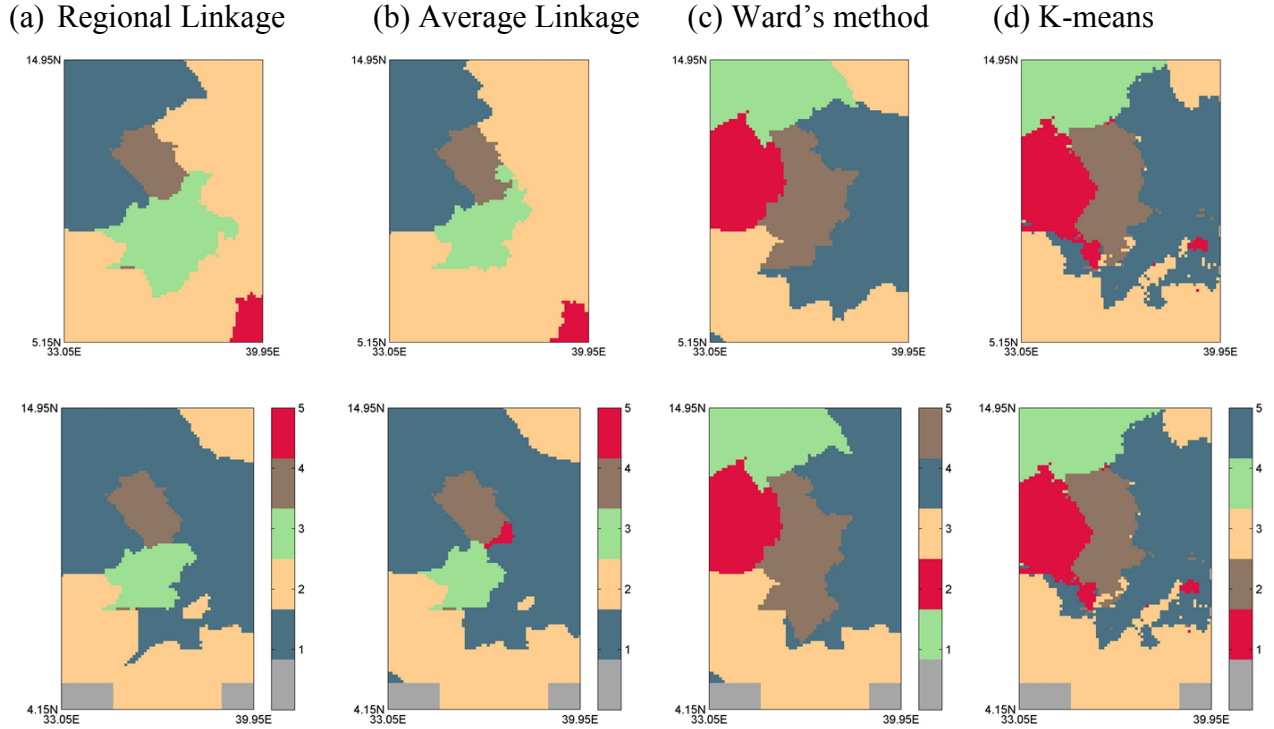


Figure 1-5: Comparative Results of hierarchical clustering specifically using (a) Regional Linkage, (b) Average Linkage and (c) Ward's method as distance criteria and non-hierarchical (d) k-means clustering over (1) the region of 5.15N – 14.95N, 33.05E – 39.95E (top row), and (2) with additional region of 4.15N – 5.15N, 34.95E – 38.75E (bottom row). All with the number of clusters $k = 5$.

Smoothing borders for k-means clusters is possible using smoothing factors, defined as the minimum number of grid-cells allowed for an isolated group to remain (i.e. grid-cells in the same cluster isolated from the main cluster). A higher smoothing factor has a higher smoothing extent. If an isolated group has fewer grid-cells than the smoothing factor (e.g. 25 grid-cells), these grid-cells will be absorbed into its adjacent cluster. If more than one adjacent cluster exists, the cluster exhibiting the higher inter-correlation with the group will be selected (Table 1-2). Smoothing with factors of 5, 10, 15 and 25 grid-cell minimums are evaluated (Figure 1-6). Even under the highest smoothing factor of 25, k-means clustering still produces the lowest WSS among all the clustering methods tested here (Table 1-1).

Table 1-1: WSS computed for clustering results in Figure 1-5

WSS*10 ⁴	Regional Linkage	Average Linkage	Ward's Method	K-means Clustering	K-means Clustering*
Data Region (1)	10.5	10.2	9.48	8.99	-
Data Region (2)	11.4	11.4	10.0	9.51	9.53

*the k-means clustering results after smoothing with a factor of 25

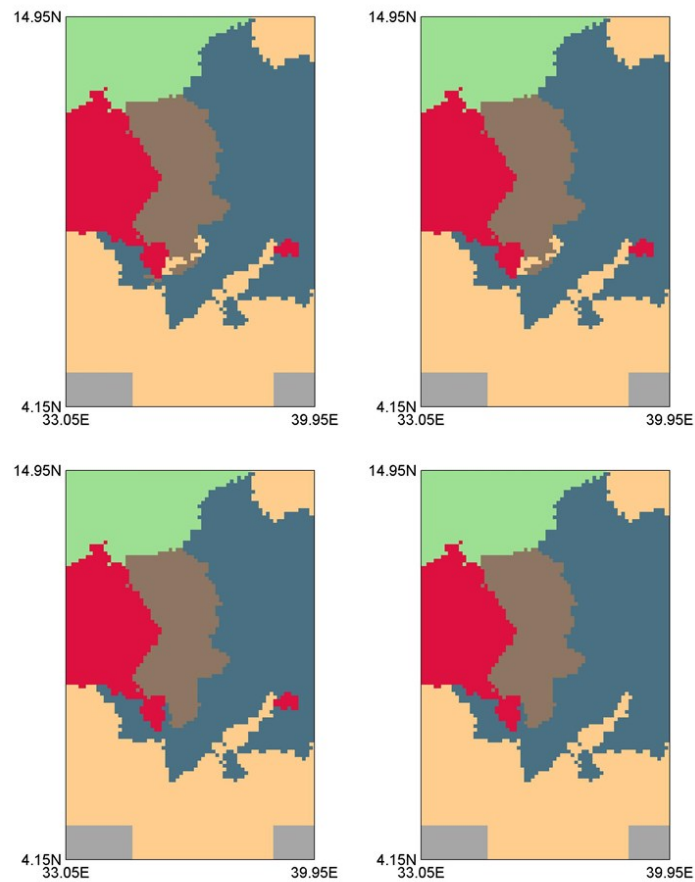


Figure 1-6: K-means clustering map under smoothing factor of 5 (top left), 10 (top right), 15 (bottom left) and 25 (bottom right)

Intra- and inter-correlations may also be compared between techniques (Table 1-2). The inter-correlation is the correlation between any two centroids (i.e. the average time series of each cluster). The intra-correlation is defined as the average correlation between the time series in each grid-cell in one cluster and the centroid of that cluster. As expected,

relatively high intra-correlations and low inter-correlations are obtained for all four clustering methods. At the 0.01 significance level, regional and average linkage produce lower inter-correlations (all lower than 0.5) than Ward's method (one surpassing 0.6) and k-means (four exceed 0.5). For the intra-correlations, k-means produces the highest overall average, all ranging between 0.64 - 0.84. The average linkage approach also produces fairly stable (consistent) intra-correlations, albeit lower overall than k-means. The Ward's method produces both the highest (0.85) and lowest (0.55) intra-correlation amongst all clustering methods; the lowest one is even less than some of its inter-correlation values. Regional linkage yields a perfect correlation of 1 within cluster 5 because a single grid-cell constitutes this cluster. This is physically unrealistic yet statistically correct, implying the over-sensitive response of the regional linkage method. Depending on the application, intra-correlation may be favored over inter-correlation, implying that homogeneous regions are more valued than independent regions.

Overall, non-hierarchical clustering tends to outperform hierarchical clustering for this particular application, given its low WSS, relatively high yet reasonable intra-correlations and acceptably low inter-correlations, and most importantly, its flexibility to produce much more stable cluster delineations. Therefore, the following sections considering the selection of k and a sensitivity analysis are restricted to the non-hierarchical (k-means) clustering technique only.

Table 1-2: Inter-correlation and Intra-correlation

(a) Regional Linkage

Cluster Index		1	2	3	4	5
Inter-correlation	1	1	0.4193	0.4453	0.3278	0.2953
	2	0.4193	1	0.2778	0.2983	0.0802
	3	0.4453	0.2778	1	0.3146	0.4306
	4	0.3278	0.2983	0.3146	1	0.013
	5	0.2953	0.0802	0.4306	0.0130	1
Intra-correlation		0.6353	0.6903	0.6542	0.7747	1.0000

(b) Average Linkage

Cluster Index		1	2	3	4	5
Inter-correlation	1	1	0.4154	0.4131	0.3126	0.2875
	2	0.4154	1	0.289	0.2945	0.2623
	3	0.4131	0.289	1	0.3275	0.388
	4	0.3126	0.2945	0.3275	1	0.2636
	5	0.2875	0.2623	0.388	0.2636	1
Intra-correlation		0.6299	0.6938	0.6582	0.7962	0.8267

(c) Ward's Method

Cluster Index		1	2	3	4	5
Inter-correlation	1	1	0.6143	0.1421	0.4958	0.4046
	2	0.6143	1	0.0852	0.4483	0.3685
	3	0.1421	0.0852	1	0.5784	0.4202
	4	0.4958	0.4483	0.5784	1	0.5156
	5	0.4046	0.3685	0.4202	0.5156	1
Intra-correlation		0.8141	0.8533	0.7162	0.6821	0.5523

(d) K-means Clustering

Cluster Index		1	2	3	4	5
Inter-correlation	1	1	0.3232	0.0677	0.5725	0.4921
	2	0.3232	1	0.3213	0.3572	0.3300
	3	0.0677	0.3213	1	0.1314	0.5207
	4	0.5725	0.3572	0.1314	1	0.4573
	5	0.4921	0.3300	0.5207	0.4573	1
Intra-correlation		0.8296	0.6430	0.7114	0.8366	0.6979

1.4 Selection of k for Non-hierarchical Cluster Analysis

As previously discussed, delineation involving the selection of k is typically performed subjectively. For low-density station-based data, this may be appropriate, however for high-resolution gridded data, an objective, automated process is appealing. Here we discuss numerous methods and their propensity for both objective selection and subjective evaluation, including the Elbow method, and Gap statistic, as well as visualization of corresponding cluster maps. New methods, including the “difference in minWSS” and “difference in difference”, are also proposed and developed to facilitate objective selection of k . Note that prediction performance is isolated from the evaluation of k intentionally, as we believe the approach adopted here – with selection based on regionalization and not prediction – while conservative, does provide a more objective and realistic expectation of prediction skill.

Previous studies addressing homogeneous June-September precipitation clusters for Ethiopia prescribe anywhere from 4 to 8 clusters; Ethiopia’s NMA officially divides Ethiopia into 8 clusters. These are all determined based on station-level data. Using the newly available high-resolution gridded dataset, but cognizant of previous work, an upper limit of 10 clusters is considered.

1.4.1 Elbow Method and “Difference in minWSS”

The Elbow method measures how the WSS decreases with increasing number of clusters. If by adding one additional cluster, the WSS improvement slows, as compared to the

previous cluster addition, an elbow will form (graphically). If this point is distinct, the optimal number of clusters selected should then be the elbow point. As previously mentioned, the common Elbow method is not suitable in this case, given that no distinct elbow is established, due to the large number of grid-cells considered yet relatively small number of clusters desired (Figure 1-4). Similarly, using AIC, BIC or GCV (Equation 5, 6 and 7; Craven and Wahba, 1979; Manning et al., 2008a) as alternative criteria only modify the curve negligibly since the large number of grid-cells produces a high value of WSS, relative to which a higher number of clusters is not strongly penalized (Figure 1-7). Therefore, an Elbow method still fails to identify the optimal number of clusters below 15. The equations for calculating AIC, BIC and GCV are listed here:

$$AIC_K = \min WSS_K + 2 \times M \times K \dots\dots\dots \text{(Equation 1-5)}$$

$$BIC_K = \min WSS_K + \log(N) \times M \times K \dots\dots\dots \text{(Equation 1-6)}$$

$$GCV_K = \frac{\min WSS_K}{\left(1 - \frac{M \times K}{N}\right)^2} \dots\dots\dots \text{(Equation 1-7)}$$

where minWSS is the minimum WSS obtained from 1000 iterative k-means clustering processes with randomly selected initial states. K is the number of clusters; note that K is a variable here. M is the number of variables or the component-wise dimensions of the data; in this case, M is the number of years. N is the total grid-cells number.

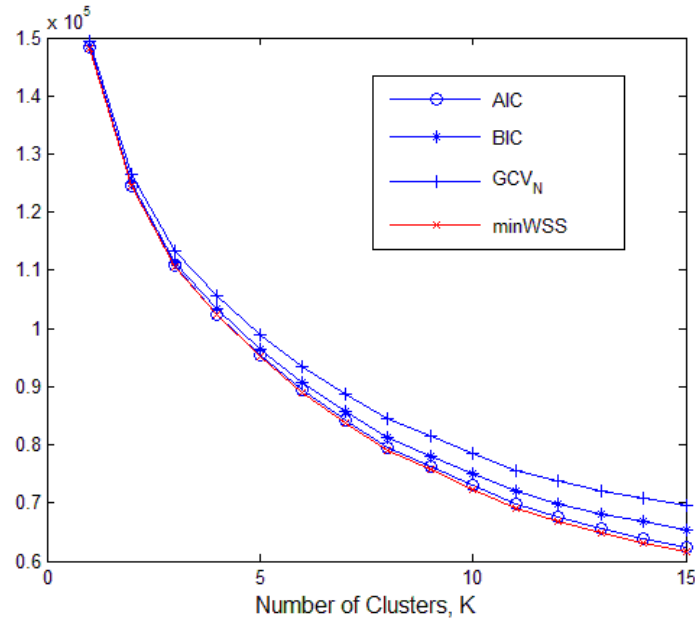


Figure 1-7: Minimum WSS, AIC, BIC and GCV given different number of clusters k based on k-means clustering results on JJAS seasonal total precipitation over the complete study region.

To extend the Elbow method, the difference in minWSS from $k-1$ to k may be calculated and evaluated (Figure 1-8). In contrast to the very smooth minWSS curve from the standard Elbow approach, the difference in minWSS shows apparent elbows (downward elbows at $k = 4$ and $k = 9$, and upward elbows at $k = 8$, through $k = 10$), indicating a significant change in the slope. In contrast to the typical Elbow method, where the reduction of the minWSS, or simply the improvement, is regarded as the base for evaluation, the “difference in minWSS” scrutinizes rates of improvement to locate the optimal k , such that subtle changes in the “smooth” minWSS curve can be captured. In this case, the difference in minWSS, that is, the decrease of WSS from $k-1$ to k , is large but initially decelerates sharply, as is typically expected for complex climate datasets, and then the deceleration becomes noticeably more gradual and nearly consistent from $k = 4$ to $k = 8$. However, from $k = 8$ to $k = 9$ the negative rate of improvement suddenly

intensifies, indicating a faster decrease of the rate of improvement – an undesirable situation relative to the previous reductions; thus, an effective selection is reached at $k = 8$. Another marginal selection of $k = 2$ may also be considered as one of the potential candidates for optimal k . This new extended method can effectively identify the best number of clusters particularly when the desired number is relatively small compared with the large number of objects processed (e.g. grid-cells), improving upon the typical Elbow approach.

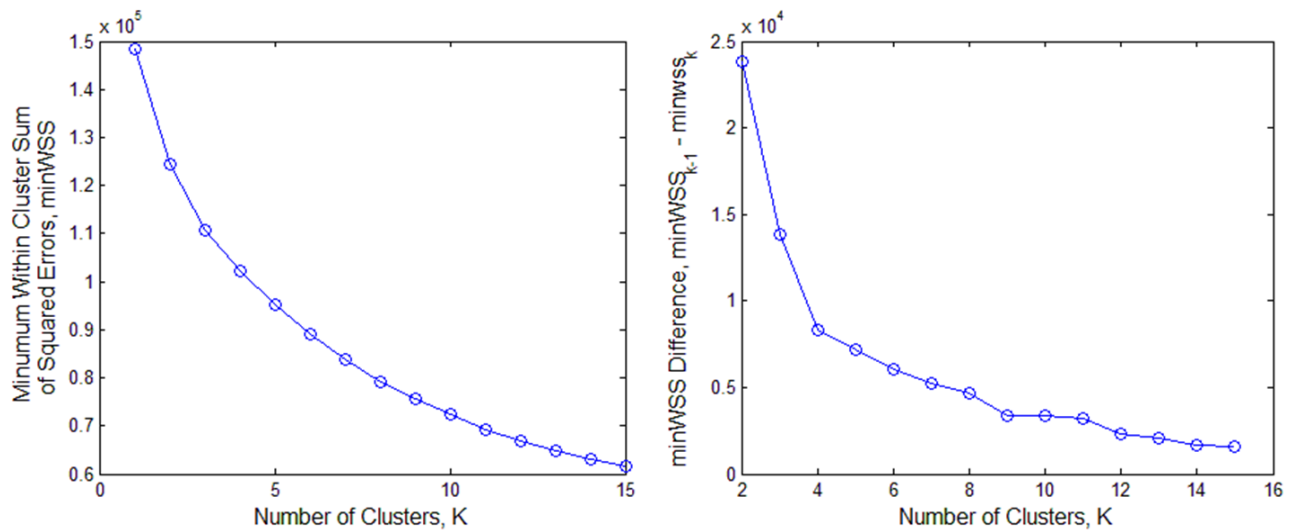


Figure 1-8: The minWSS and difference in minWSS from $k-1$ to k .

1.4.2 Gap Statistic and “Difference in Difference”

For a gap statistic approach (Tibshirani et al., 2001), the WSS from a cluster analysis on a randomly simulated (reference) data set having the same dimensions as the original data set is compared to the WSS from a cluster analysis on the original data set. Intuitively, if clustering of the original data set provides a similar WSS to the randomly simulated data

set, which should not have any cluster characteristics given its random nature, clustering of the original dataset is deemed inappropriate. A large difference between the WSS from the random and original datasets is preferred for the selection of k . This difference is the so-called “gap”.

The gap statistic algorithm includes:

- (1) generate a reference data set
- (2) perform cluster analysis with varying k
- (3) compute the corresponding WSS_K^*
- (4) iterate steps (1) - (3) B times.

A reference WSS in logarithmic form is then calculated as the expectation of all WSS_K^* for each k (Equation 1-8). The gap is the difference between the WSS_K from the original dataset and WSS_K^* from the reference dataset, both in logarithmic form (Equation 1-9):

$$E\{\log WSS_K^*\} = \frac{1}{B} \sum_{b=1}^B \log WSS_{K,b}^* \dots\dots\dots \text{(Equation 1-8)}$$

$$Gap_K = E\{\log WSS_K^*\} - \log WSS_K \dots\dots\dots \text{(Equation 1-9)}$$

The standard deviation sd_K and the simulation error se_K based on the reference dataset are also required (Equation 1-10 and 1-11):

$$sd_K = \left\{ \frac{1}{B} \sum_{b=1}^B \left[\log WSS_{K,b}^* - E\{\log WSS_K^*\} \right]^2 \right\}^{1/2} \dots\dots \text{(Equation 1-10)}$$

$$se_K = \sqrt{1 + 1/B} \cdot sd_K \dots\dots\dots \text{(Equation 1-11)}$$

The optimal number of clusters is then defined as the smallest k such that

$$Gap_K - (Gap_{K+1} - se_{K+1}) \geq 0 \dots\dots\dots \text{(Equation 1-12)}$$

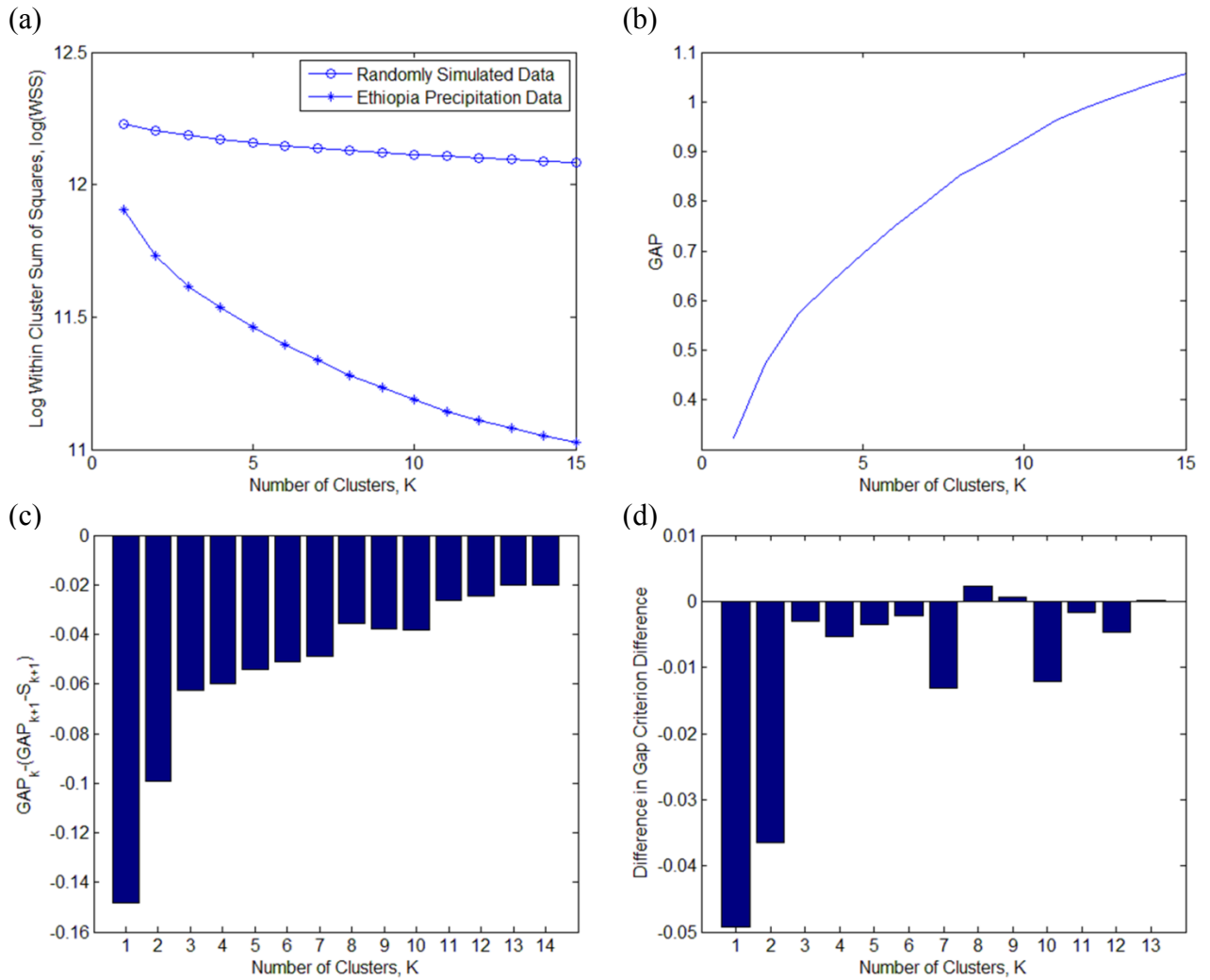


Figure 1-9: Gap statistic and “difference in difference” results.

The difference in gaps considering simulation errors is the “gap criteria” (Equation 1-12).

The gap statistic is applied with $B = 100$ trials and $k \leq 15$. In this case, results show that as k increases, so does the gap (Figure 1-9a, b), and the gap criteria has not reached a nonnegative value prior to $k = 15$ (Figure 1-9c), indicating an optimal k greater than 15. However, as previously discussed, having more than 15 clusters is not preferred in this case. Thus an additional step is added by computing the difference in gap criteria, called the “difference in difference” (Figure 1-9d), to measure the “speed vector” rather than the

“state” of reaching the nonnegative value along the “time step” k . This serves as the secondary criterion for situations when the optimal number of clusters exceeds the desired limit using the traditional gap statistic approach. The first nonnegative “difference in difference” occurs at $k = 8$, illustrating the decline of the gap criteria from $k = 8$ to $k = 9$. In contrast, a relatively large improvement occurs between $k = 7$ and $k = 8$. Therefore, $k = 8$ is a suitable number of clusters, falling into our desired range of k , which is consistent with the “difference in minWSS” approach, further supporting the selection of $k = 8$.

1.4.3 Visualization of Cluster Maps

Visualization of cluster maps for different k is a more direct way to select an appropriate number of clusters, albeit subjective. Nonetheless, it is a useful approach to confirm results from the objective measures described above. Cluster maps ($k = 2$ to $k = 9$; Figure 1-10) indicate the stability of certain clusters even as k changes; for example, the central zones remain relatively constant from $k = 4$ to $k = 5$ and from $k = 6$ to $k = 9$. For $k = 2$ through $k = 5$, a relatively low number of clusters tends to produce cluster partitioning; the northeastern and southern regions typically belong to the same cluster but are geographically separated. It is not until $k = 7$ that the two regions are completely assigned to separate clusters. Given that clusters represent homogeneous precipitation patterns, it is unlikely (however possible) that disjointed clusters make physical sense; it is instead preferred that such clusters be split into distinct independent clusters. Thus for this dataset, $k \geq 7$ is preferred. Comparing $k = 8$ and $k = 9$, cluster boundaries appear

very similar, however the former illustrates a cleaner delineation, minimizing the jumble of clusters in the southeast portion of the dataset evident for $k = 9$.

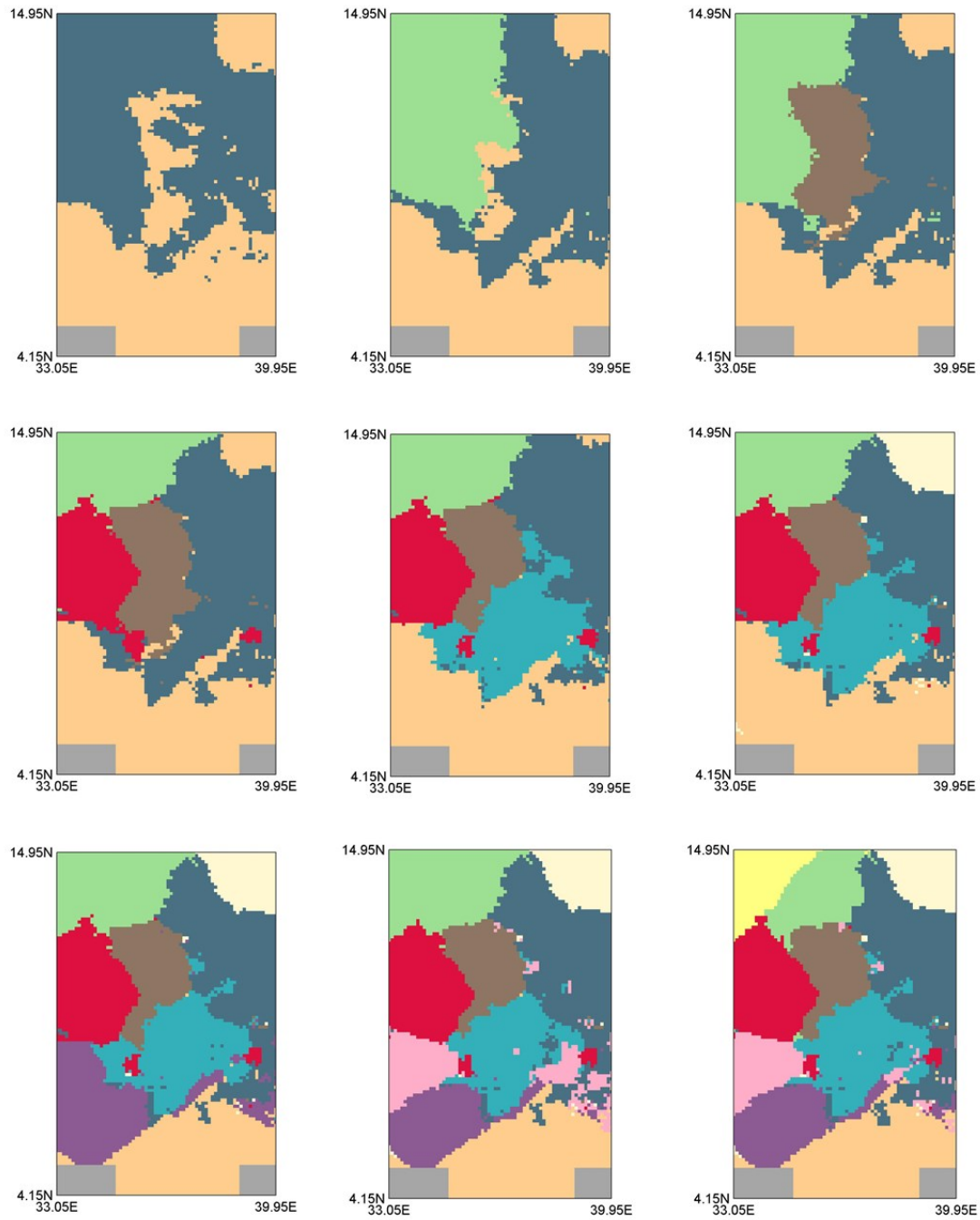


Figure 1-10: K-means clustering maps given different number of clusters k ranging from 2 to 10

1.4.4 Sensitivity Analysis using Reduced Time series

Although non-hierarchical clustering is relatively immune to extending spatial data, it can be subject to data length. Three shortened time series are compared with the original time series of 29 years by dropping 5 years: a) the first 5 years, b) the last 5 years, and c) the driest 5 years. The driest 5 years are defined based on the average JJAS seasonal total precipitation over the complete study region (Figure 1-11) and include 1984, 1987, 1997, 2002 and 2011. Interestingly, for $k = 7, 8$ and 9 , dropping the driest 5 years has the least influence of the three shortened time series as compared with the original outcomes (Figure 1-12). This is predominantly attributable to all grid-cells tending to behave similarly in drought years given the extensive range of consistently dry conditions throughout the region. Thus the lack of differentiation spatially contributes very little to the cluster analysis. On the contrary, dropping moderate years affects clustering outcomes to a larger extent given spatial variability. Thus dropping the first 5 years of data, which contains two extreme years but three very moderate years, and the last 5 years, which contains one extreme year and four moderate years, produces notably different cluster boundaries (Figure 1-12).

Reducing the time series length provides less variability in the cluster analysis and thus results in more fractured clusters; in other words, the remaining variability is less likely to clearly distinguish clusters or identify their homogeneity. However most cluster patterns are still recognizable, indicating robust regionalization.

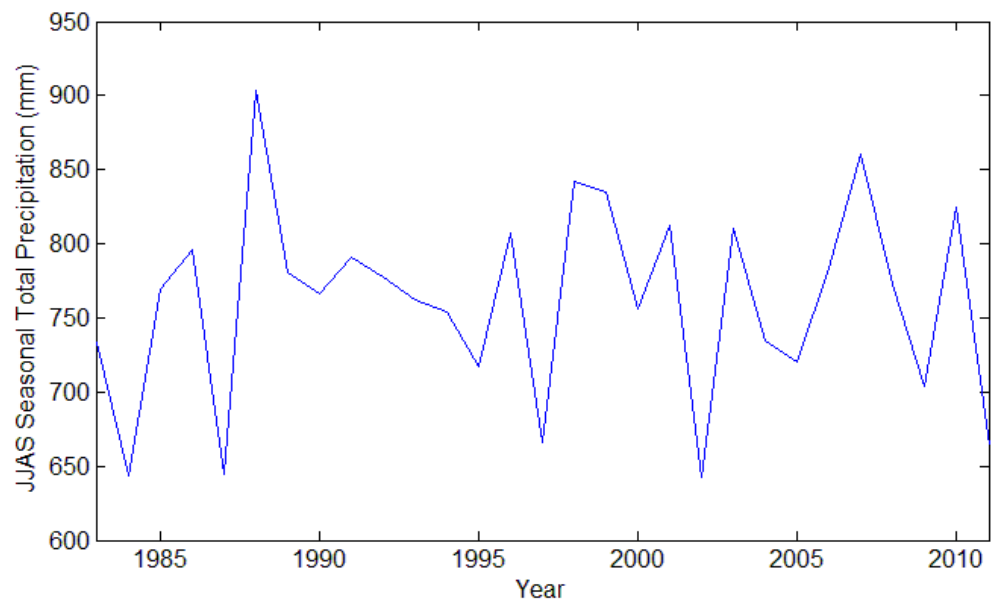
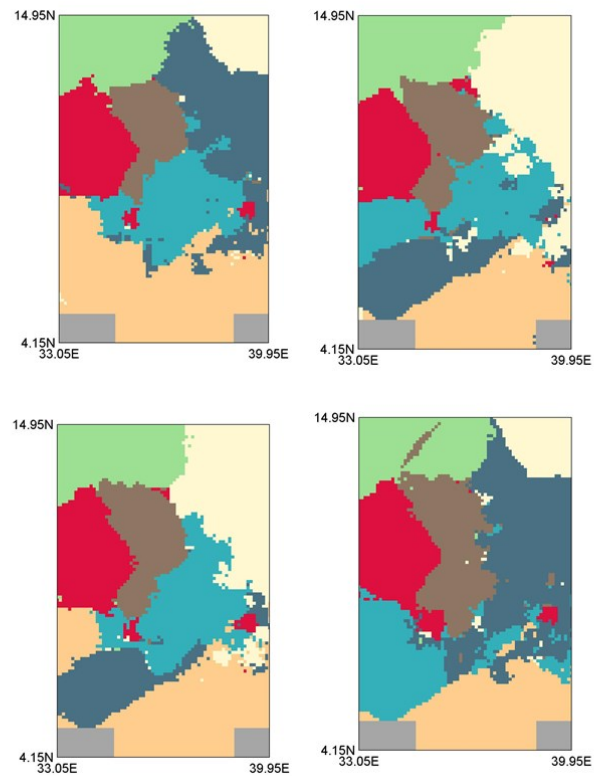
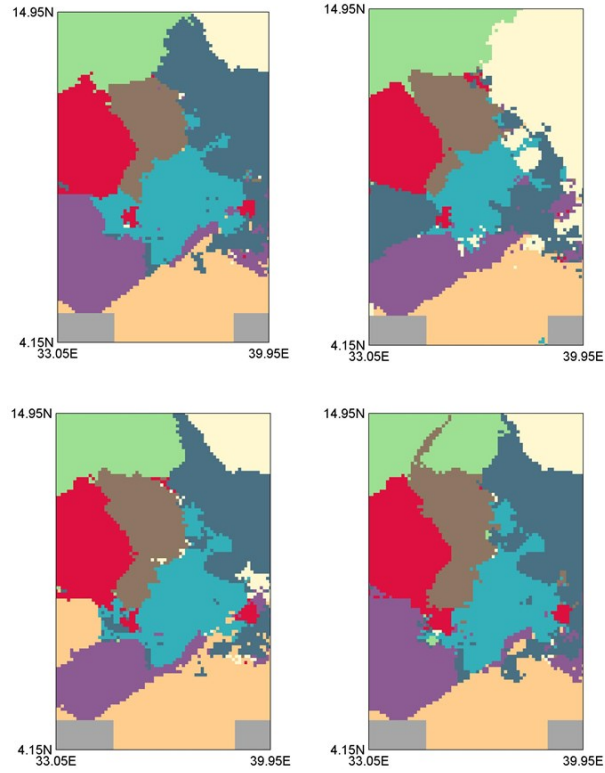


Figure 1-11: Mean time series of JJAS seasonal total precipitation over the complete study region.

(a) $k = 7$



(b) $k = 8$



(c) $k = 9$

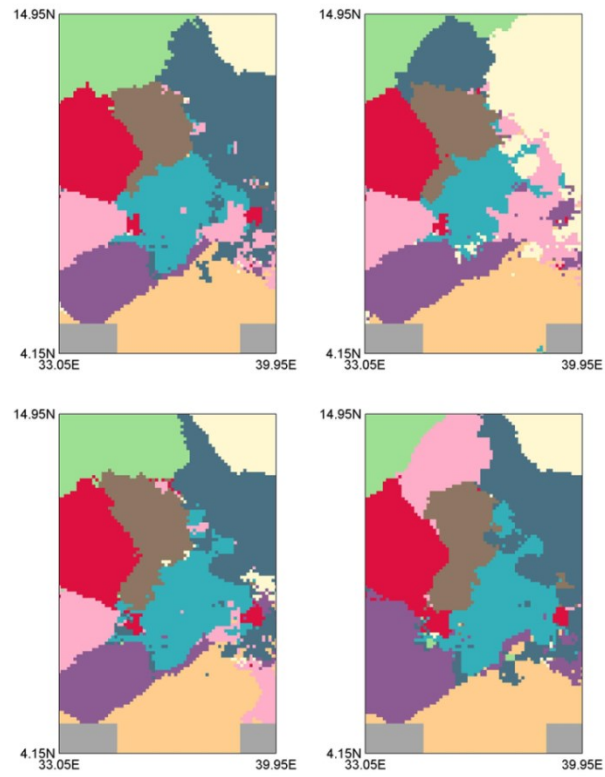


Figure 1-12: clustering maps with full time series (top left), drop 5 driest years (bottom left), drop first 5 years (top right), and drop last 5 years (bottom right) for $k = 7, 8$ and 9 .

In summary, for k-means clustering, the delineation is affected by the number of predetermined clusters and is sensitive to the length of available data and characteristics of historical years. For this Ethiopia precipitation analysis, $k = 8$ appears superior given the objective measures – both the “difference in minWSS” based on the elbow method and “difference in difference” based on the gap statistic – and the subjective measures through visualization including spatially coherent clusters and relatively smooth boundaries. For lesser numbers of clusters, single clusters fractured into multiple sub regions, which are geographically distant, is problematic and undesirable. Intra- and inter-correlations for $k = 8$ (Table 1-3) indicate an overall strong coherency within each cluster, further justifying selection of $k = 8$ as an optimal choice (Figure 1-13). Individual cluster mean time series (Figure 1-14) also indicate diversity between clusters.

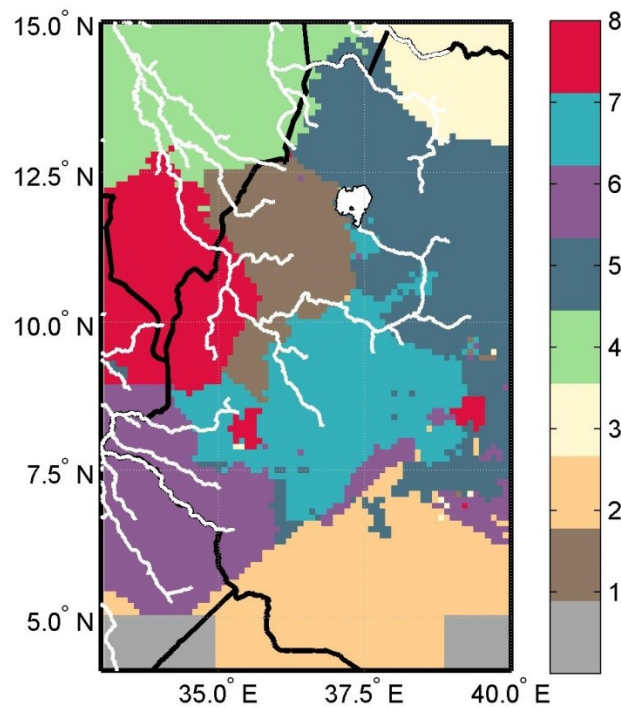
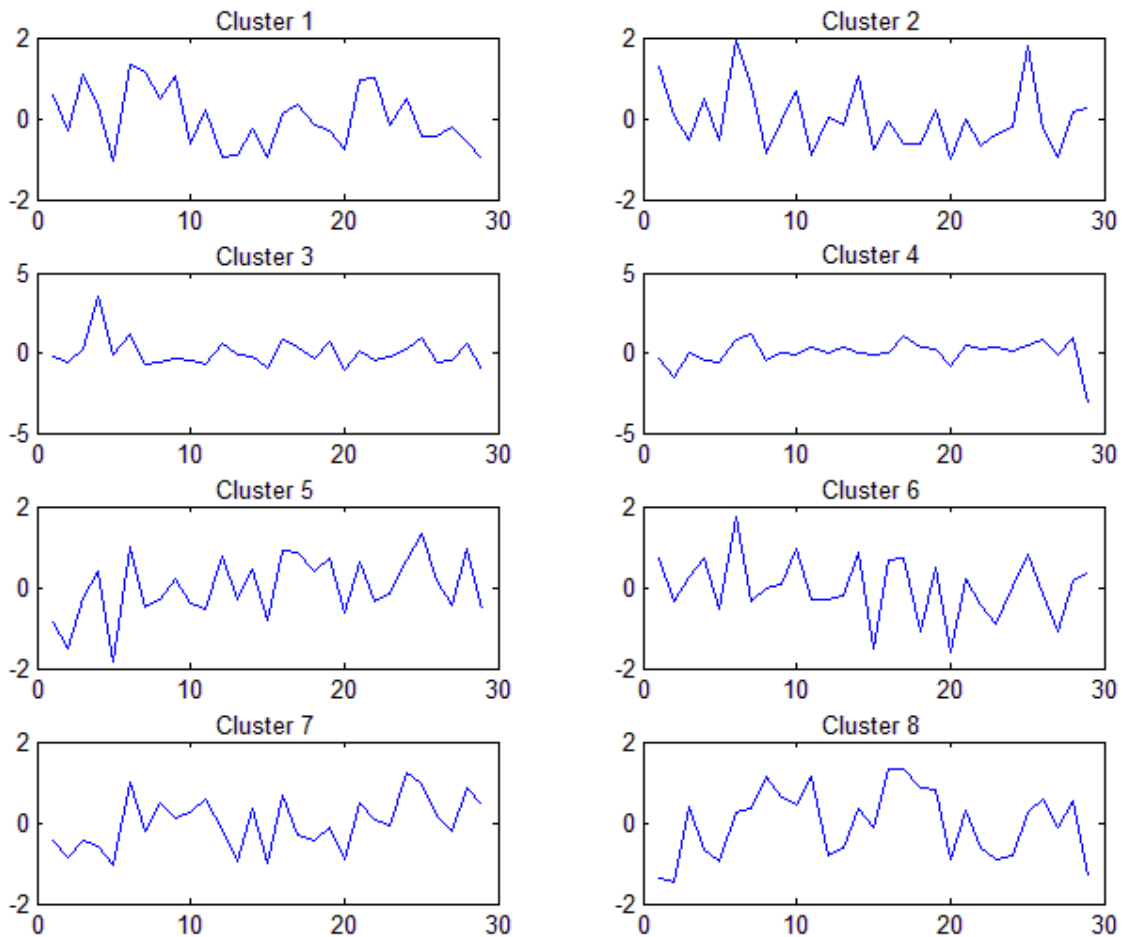


Figure 1-13: k-means clustering maps given $k = 8$ with country boundary and river profile.

Table 1-3: intra- and inter-correlation table for k-means clustering results given $k = 8$

Cluster Index		1	2	3	4	5	6	7	8
Inter-correlation	1	1	0.1861	0.2023	0.3853	0.2322	0.3566	0.3168	0.3042
	2	0.1861	1	0.4047	0.0916	0.3548	0.739	0.3625	-0.0848
	3	0.2023	0.4047	1	0.2182	0.5533	0.5497	0.1615	0.0684
	4	0.3853	0.0916	0.2182	1	0.4936	0.1171	0.2096	0.5725
	5	0.2322	0.3548	0.5533	0.4936	1	0.5002	0.6381	0.4996
	6	0.3566	0.739	0.5497	0.1171	0.5002	1	0.5289	0.2584
	7	0.3168	0.3625	0.1615	0.2096	0.6381	0.5289	1	0.3844
	8	0.3042	-0.085	0.0684	0.5725	0.4996	0.2584	0.3844	1
Intra-correlation		0.7380	0.7907	0.9051	0.8488	0.773	0.7753	0.6412	0.8544

Figure 1-14: Standardized mean time series (centroid) of JJAS seasonal total precipitation within each clustered region from 1983 to 2011 (29 years); clustering results are based on $k = 8$.

1.5 Identification of Local and Large-scale Cluster-level Precipitation Drivers

Another technique to differentiate the independence of clusters is to understand the local and large-scale drivers effecting precipitation patterns and variability. This is typically undertaken by evaluating correlations between cluster-level precipitation (JJAS seasonal total in this case) with hydroclimatic variables. From these, physical mechanisms may be explored and identified. Gridded ($2.5^{\circ} \times 2.5^{\circ}$) global NCEP/NCAR Reanalysis data (Kalnay et al., 1996) for five different climate variables selected – sea surface temperature (SST), sea level pressure (SLP), geopotential height (GH) at 500mb, surface air temperature (SAT) and outgoing longwave radiation (OLR) – is correlated with the mean time series of JJAS seasonal total precipitation from each cluster. Previous studies have identified relationships between Ethiopia's precipitation and large-scale climate variables, such as SSTs in equatorial Pacific, Indian Ocean and equatorial/southern Atlantic Ocean and SLPs near the African continent, specifically the Azores High, St. Helena High, Mascarene High and southwest Asian Monsoon (Korecha and Barnston, 2007; Segele et al., 2009; Diro et al., 2011b).

As expected, each cluster is associated with a unique set of large-scale climate patterns, although similarities exist. For example, similar signals from concurrent-season SST patterns in the equatorial Pacific region (representing ENSO) are found across all clusters but to varying degrees (Figure 1-15a). From north to south of western Ethiopia, the influence of ENSO generally decreases. Northwestern Ethiopia, particularly the region close to the Rift Valley (Cluster 5), is the most strongly negatively correlated with equatorial Pacific SST. This association weakens southward (central-western Ethiopia

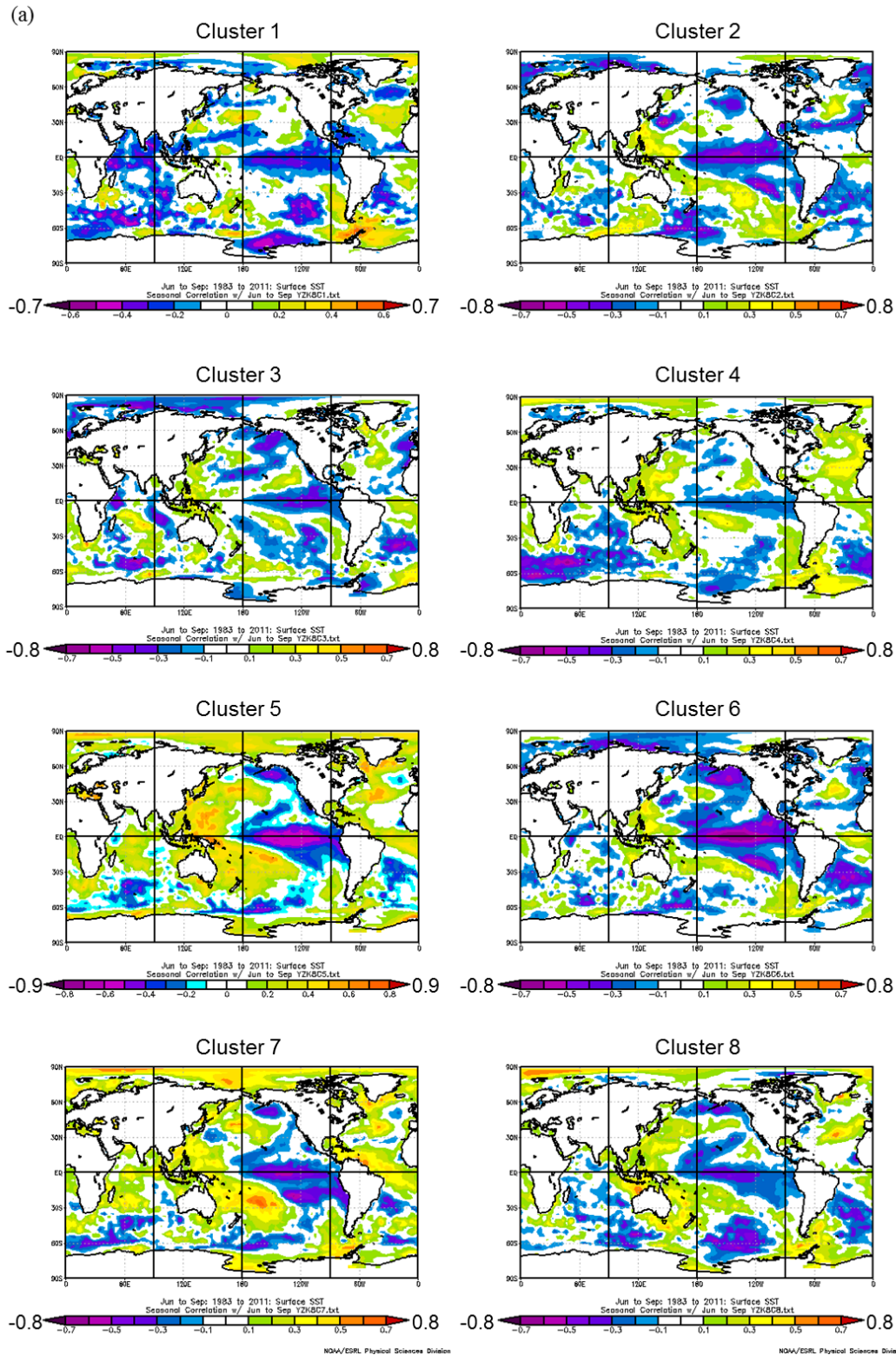
and then southwestern Ethiopia; Cluster 7, 6 and 2), indicating a smaller influence from ENSO, but still significant. All of the clustered regions illustrate a negative correlation, that is, warmer/colder equatorial Pacific SST (El Niño/La Niña) typically bring deficit/excess JJAS seasonal total precipitation to the study region. This concurs with a number of previous studies (e.g. Segele and Lamb, 2005; Korecha and Barnston, 2007; Korecha and Sorteberg, 2013). On the other hand, weak, insignificant associations with equatorial Pacific SST exist for some regions, including the Southwest corner of Sudan (Cluster 4) and its neighboring region from part of northwestern Ethiopia (Cluster 1). These specific regions, however, show stronger correlations with Indian Ocean and southern Atlantic Ocean SST, perhaps implying a more direct influence on moisture transport from these oceanic regions (Figure 1-15a). The overall influence of the Indian and southern Atlantic Oceans, as compared to the Pacific Ocean, on JJAS seasonal total precipitation in western Ethiopia is still less, based simply on the correlation with SST in each ocean.

Global correlation maps between concurrent-season SLPs and cluster-level time series illustrate diverse positive and negative patterns (Figure 1-15b). For example, the Mascarene high pressure system in the southern Indian Ocean is apparent for southwestern Ethiopia (Clusters 2 and 6). As the moisture-laden winds caused by the Mascarene High cross the equatorial Indian Ocean from south to north, they change direction from southeasterly to southwesterly due to the Coriolis force (Viste and Sorteberg, 2013b), affecting precipitation in southwestern Ethiopia. The Mascarene High is negatively correlated with JJAS precipitation in southwestern Ethiopia, and may control the extent to which the ITCZ shifts. Previous studies (e.g. Segele et al., 2009)

indicate that the Mascarene anticyclone is coupled to a weak, semipermanent surface ridge, which appears to limit the southern range of the ITCZ during the northern hemisphere summer. It is speculated that the timing of the development of this pressure system also matters. For example, an early development of the Mascarene high pushes the ITCZ to the north prior to JJAS, followed by an earlier diminish of the pressure system during JJAS (low pressure). Meantime, ITCZ moves back to the south, and brings more rainfall to the southern part of western Ethiopia (high precipitation, and thus negative correlated with low pressure at Mascarene High). Additionally, local effects from adjacent SLP to the north of Ethiopia are non-negligible for regions extending from northwestern to southwestern Ethiopia (Clusters 5, 6, 7 and 8), coinciding with previous findings identifying a moisture flow path connecting the Mediterranean, Red Sea, and Arabian Peninsula with the northern Ethiopian highlands in the summertime (Viste and Sorteberg, 2013a; Viste and Sorteberg, 2013b).

Concurrent-season SAT time series over the Sahel indicate remarkably strong negative correlations with seasonal precipitation for a large portion of northwestern and central-western Ethiopia (approximately -0.9 for northwestern Ethiopia; Figure 1-16a). It is likely due to a decrease of monsoon-related continental convergence and rainfall from Senegal to Ethiopia (Giannini et al., 2003), causing a high SAT over Sahel and low precipitation in Ethiopia. For concurrent-season OLR, the local index correlates strongly with southwestern Ethiopia (Cluster 2 and 6) (Figure 1-16b). Therefore, summertime precipitation in the southern part of the study region is likely to be more influenced by local climate variables compared with other clusters. Not surprisingly, clusters with high inter-cluster correlations (e.g. 2 and 6, 5 and 7; Table 1-3) tend to produce similar

correlation patterns and are therefore expected to be similarly affected by local and large-scale precipitation drivers.



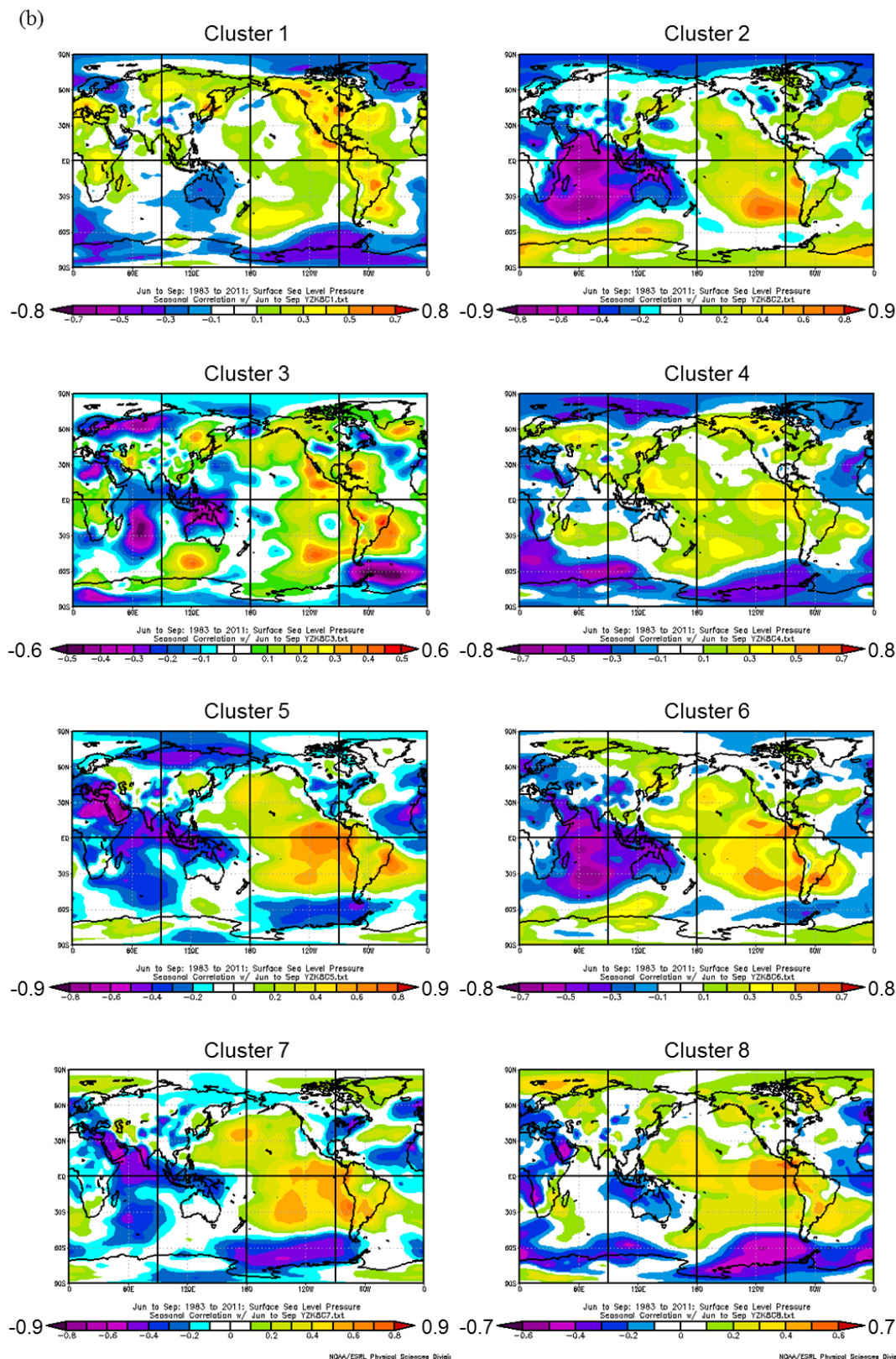


Figure 1-15: Global correlation map of (a) SST and (b) SLP correlated with within-cluster mean time series of JJAS seasonal total precipitation.

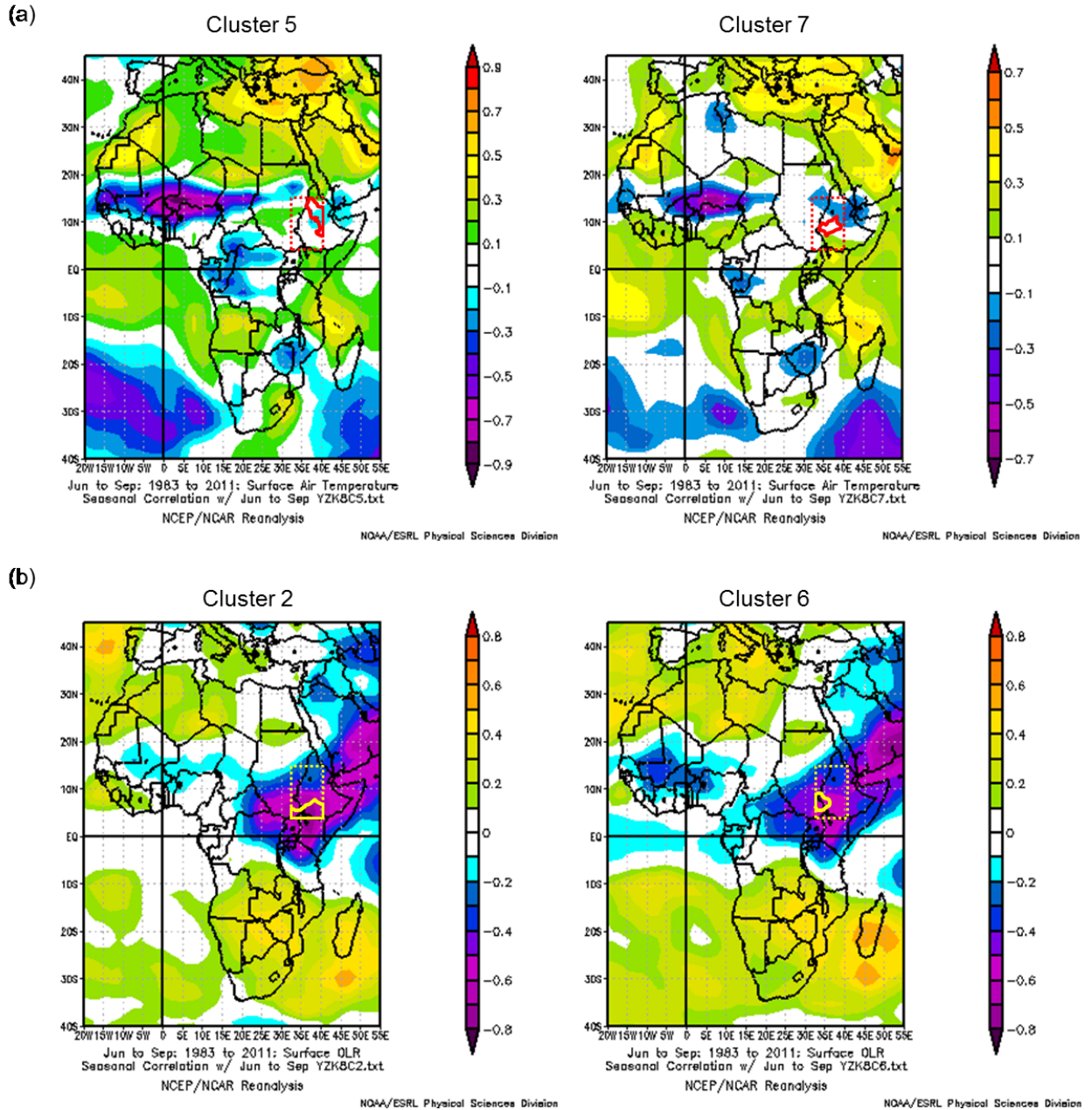


Figure 1-16: Correlation map of (a) surface air temperature, (b) outgoing longwave radiation correlated with within cluster mean time series of JJAS seasonal total precipitation for (a) Cluster 5 and 7, (b) Cluster 2 and 6, centered at Africa, with dashed square indicating the study region and solid line circling the corresponding cluster region.

1.6 Predictions

One motivation for developing and evaluating clusters is to produce homogeneous regions for which precipitation may be uniformly predicted and subsequently applied to agricultural planning. To demonstrate, a simple principal component linear regression model (Equation 1-13a, b) is applied to predict the JJAS seasonal precipitation total for each cluster through a drop-one cross-validation approach. This is arguably a simplified prediction technique, but sufficient for demonstration purposes here. Predictor variables, including SST, SLP, GH, and SAT, from previous month (May) over highly correlated and physically justifiable regions (Table 1-4) are selected. PCA is then performed on all selected climate variables to remove multicollinearity and reduce the number of predictors. Consequently, the top two PCs for each cluster explain approximately 52% - 82% of the total variance and are applied as predictors in the regression framework. The model expressed through equations is shown below:

$$Y_{-i} = \tilde{\alpha}_i + \tilde{\beta}_i PC1_{-i} + \tilde{\gamma}_i PC2_{-i} + \varepsilon \dots\dots\dots \text{(Equation 1-13a)}$$

$$\hat{Y}_i = \tilde{\alpha}_i + \tilde{\beta}_i \widehat{PC1}_i + \tilde{\gamma}_i \widehat{PC2}_i \dots\dots\dots \text{(Equation 1-13b)}$$

where for each cluster, Y_{-i} is the within cluster JJAS seasonal total precipitation with year i left out. Similarly $PC1_{-i}$ and $PC2_{-i}$ are the top two PCs based on the PCA of selected predictors with year i left out. $\hat{Y}_i$ is the modeled precipitation for year i based on the coefficients ($\tilde{\alpha}_i$, $\tilde{\beta}_i$, and $\tilde{\gamma}_i$) estimated from the linear regression (Equation 1-13a) and its own reconstructed PCs ($\widehat{PC1}_i$ and $\widehat{PC2}_i$; Equation 1-13b). Note that each cluster has a unique set of model inputs.

Cross-validated predictions, including a 95% confidence interval conditioned on model errors, are generally quite skillfull and closely mimic observed year-to-year variability (Figure 1-17). Pearson correlations between mean predictions and observations are greater than 0.7 for all clusters. Additionally, the median Rank Probability Skill Score, a categorical measure, ranges from 19% - 75% for all clusters, indicating significantly more skill than a climatological (historical averages) prediction (Table 1-5).

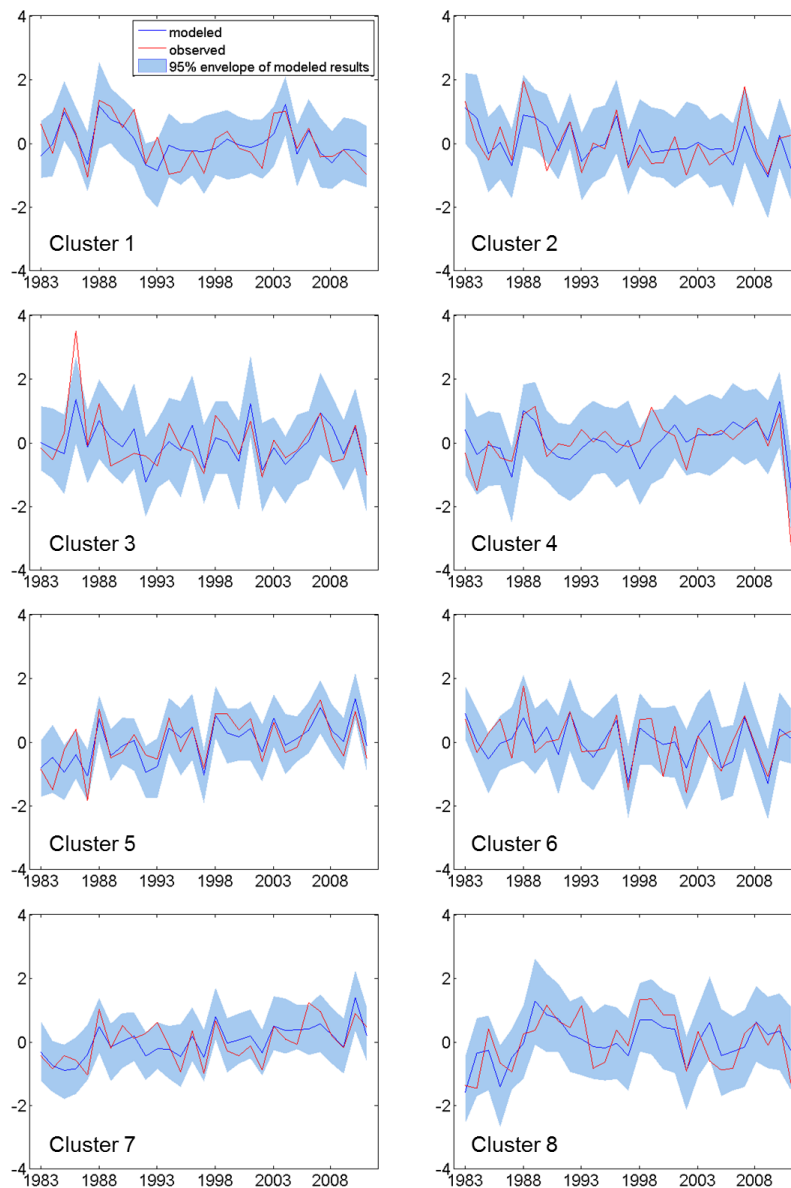


Figure 1-17: Modeled results using the first two PCs of selected climate variables as predictors, with drop-one-year cross-validation.

Table 1-4: Climate Variables (C.V.) in May over different regions for each cluster (C1 ~ C8) and region as a whole (non-cluster) used as predictors, with corresponding correlation between the climate variable averaged over the region and the cluster-level JJAS seasonal total precipitation time-series shown (only cells with correlation values shown are used as pre-predictors).

C.V.	SST				SLP						GH at 500mb						SAT	# of pre-predictors
Region	EP	NI	SI	E/SA	LO	EP	AH	SH	MH	AM	LO	EP	AH	SH	MH	AM	LO	
C1	-0.46	-0.46	-0.55	-0.48			0.45	0.45	-0.51			0.52					0.50	9
C2		-0.43	-0.51	-0.43					0.58			0.55			0.50			6
C3			-0.58	-0.59				-0.50					0.57					4
C4			-0.60							-0.58			0.49					3
C5	-0.52	0.52	-0.54	0.59	-0.50	0.61						0.67	0.67	0.54	0.53		0.67	11
C6		0.56		-0.51		0.64						0.66		-0.51				5
C7		0.63	-0.59	0.65				0.44					0.65	0.65			0.44	7
C8			-0.44	0.53	-0.46		0.55						0.63				0.48	6
Non-cluster	-0.47		-0.47	-0.52		0.47				-0.41			0.54	0.58	0.52		0.52	9

Note: EP - equatorial Pacific region, NI - North Indian Ocean, SI - South Indian Ocean, E/SA - equatorial/southern Atlantic Ocean
LO - local region, AH - Azores High, SH - St Helena High, MH - Mascarene High, AM - southwest Asian Monsoon.

Table 1-5: Pearson correlation coefficients (Corr.) and median RPSS values corresponding to the cross-validated (CV) modeled results compared with observations in Figure 1-17

C index	C1	C2	C3	C4	C5	C6	C7	C8
Corr.CV	0.741	0.695	0.711	0.683	0.838	0.744	0.751	0.699
Median RPSS (%)	45.23	26.04	36.16	19.82	74.30	5.44	51.91	48.21

1.7 Conclusions and Discussion

In this chapter, we evaluate various regionalization methods for objective delineation and define a number of approaches for optimally selecting an appropriate number of clusters, with a demonstration on western Ethiopia's summertime precipitation. Given a high-resolution gridded precipitation dataset, objective, automated clustering and delineation is possible. The preference of hierarchical versus non-hierarchical (k-means) clustering is typically case-specific; in this case k-means clustering outperforms hierarchical clustering, particularly considering its stable cluster patterns.

Various statistical methods are available objectively to define the optimal number of clusters; however these approaches fail for datasets with large numbers of grid-cells and a desire for a relatively small number of clusters. Therefore, a "difference in minWSS" extension to the Elbow method and a "difference in difference" extension to the Gap statistic method are developed to objectively define the optimal number of clusters within the desired range. Visualization of cluster maps – a subjective tool – can reinforce objective outputs from these newly developed methods.

Only a few studies have explored regionalization over Ethiopia and to our knowledge all use station data and therefore require subjective delineation. The k-means clustering result given the optimal selection of $k = 8$ tends to generally agree with Diro et al.'s (2009) grouping of stations for summertime precipitation, with almost all stations still falling into equivalent clusters. Delineation of the boundaries differs, however, given the subjectivity of drawing boundaries conditioned on sparsely located stations. Other studies (e.g. Eklundh and Pilesjö, 1990; Gissila et

al., 2004; Korecha and Sorteberg, 2013) differ in both the assignment of stations to clusters and the delineation of homogenous regions, given different datasets, clustering techniques and, most critically, the variability in precipitation considered. In other studies the overall interannual variability and month-to-month seasonality is evaluated, whereas the interannual variability of JJAS seasonal total precipitation is isolated in this chapter for regionalization, partially agreeing with Diro et al. (2009), focusing purely on the main rainy and agricultural season.

The regionalization techniques and evaluation metrics developed in this chapter can also be generally applied to other hydroclimatic datasets, serving different purposes. For studies focused exclusively on Ethiopia, the country mask can be applied; however given the robust cluster patterns produced by k-means, the clustering results would be similar (not shown). Regionalization can also be applied on subset months such as JJ and AS, when the physical mechanism of precipitation patterns migrating over the season is desired or a sub-month precipitation prediction is the subsequent goal.

It is unclear how future climate changes may affect regionalization. If it imposes regionally specific changes, clusters may eventually shift and reorient, however if climate change influences the overall region consistently, then shifting of cluster boundaries may be minor.

Chapter 2 Does Objective Cluster Analysis Serve as a Useful Precursor to Seasonal Precipitation Prediction at Local Scale? Application to Western Ethiopia

2.1 Primer on Prediction Models and Cluster Analysis

Seasonal precipitation prediction aims to provide potentially actionable information to guide planning and management of various sectoral activities. Two types of models are commonly used for seasonal precipitation prediction: statistical and dynamical. Dynamical models, such as general circulation models (GCMs), include complex physical climate processes, while statistical models are purely data-driven, relating observations and hydroclimatic variables directly.

While both modeling approaches have produced skillful seasonal predictions for a variety of applications (e.g. Barrett, 1993; Hammer et al., 2000; Shukla et al., 2016), each has noteworthy drawbacks. Dynamical models often require a significant amount of time to build and parameterize, whereas statistical models require considerably fewer resources (e.g. Mutai et al., 1998; Gissila et al., 2004; Block and Rajagopalan, 2007; Diro et al., 2008; Diro et al., 2011b; Block and Goddard, 2012). Dynamical models also suffer from their high sensitivity to initial uncertain conditions, particularly given a long lead time. Consequently, a number of simulations are typically produced, each with unique initial conditions, to provide a range of possible outcomes (e.g. Roeckner et al., 1996; Anderson et al., 2007). Furthermore, the outputs from dynamical models often require additional bias correction, typically using statistical methods, to better match observations (e.g. Ines and Hansen, 2006; Block et al., 2009; Teutschbein and Seibert, 2012). Statistical models, on the other hand, are highly dependent on substantial high-quality historical data to capture hydroclimatic patterns and signals, particularly extreme

conditions, which is often not available. Additionally, statistical models are often linear by construction, and may not well capture non-linear complex interactions and feedbacks. The physical nature of dynamical models, however, allows for prediction under non-stationary conditions, and also when insufficient historical data is available, whereas statistical models, by construction, typically rely on stationary relationships (Schepen et al., 2012).

Given the advantageous features of both model types for seasonal prediction, many studies have explored the combination of statistical and dynamical model outputs (e.g. Coelho et al., 2004; Block and Goddard, 2012; Schepen et al., 2012). In general, the combined predictions are typically superior to individual models, however this is not always the case, and is dependent on location, predicted seasons, lead time, and comparable model skill. (e.g. Metzger et al., 2004).

The spatial extent selected for statistical seasonal prediction is critical. It is not uncommon to simply assume homogeneity in precipitation across an entire study region, however this limits addressing potential spatial variability. While this may be suitable for very broad regional planning, it is often ineffectual for operational and local-level decisions, particularly for regions with high spatial variability. This has led to the need for defining sub-regional scale homogeneous regions. Defining these homogeneous regions, however, is a non-trivial process. There are a variety of methods, including simply comparing annual cycles (e.g. unimodal and bimodal distributions in precipitation) between stations (or grid-cells), comparing station correlations with all-region averages, or applying empirical orthogonal functions (EOF) and various clustering techniques to define homogeneous sub-regions, and others of increasing complexity (e.g. Parthasarathy et al., 1993; Mason, 1998; Landman and Mason, 1999; Gissila et al., 2004; Diro et al., 2008; Diro et al., 2011b; Singh et al., 2012). In addition, delineation of the

sub-region size is also important to consider. Smaller sized homogeneous sub-regions do not necessarily lead to improved predictions, as the noise at overly small scales can dominate any real signals representing spatial coherency of precipitation. For additional discussion regarding defining homogeneous sub-regions and cluster analysis, the reader is referred to Chapter 1 of this dissertation and Badr et al. (2015).

Operational precipitation predictions in Ethiopia have been issued by its National Meteorological Agency (NMA) since 1987 using an analog methodology (i.e. locating a similar climate scenario in the past – an analog – to predict future conditions), however this approach has produced only marginally skillful outcomes (Korecha and Sorteberg, 2013). For NMA's prediction, the country is divided into eight homogeneous regions for which NMA produces independent predictions. Similarly, others have also addressed seasonal prediction in Ethiopia contingent on both temporal and spatial precipitation patterns. Gissila et al. (2004) divide Ethiopia into four regions conditioned by the seasonal cycle and interannual variability coherence prior to prediction, while Diro et al. (2009) apply a similar approach but with dynamic cluster boundaries, allowing for different delineations for each rainy season. Segele et al. (2015) consider statistical precipitation predictions across Ethiopia as a whole, as well as for northeastern Ethiopia and at two Ethiopian cities. Block and Rajagopalan (2007) predict the average summertime (June – September, or JJAS) precipitation over the upper Blue Nile basin – a region they claim is homogenous at interannual time scales. Korecha and Barnston (2007) select an all-Ethiopia average precipitation index to characterize predictability broadly, with minimal attention to operational-level predictions. All of these studies focus on predicting regional average precipitation based on subjective clustering methods applying a limited number of stations or coarsely gridded data; no local predictions at a finer spatial scale are explored.

This study moves forward by exploring local-level seasonal precipitation prediction through the use of sub-regional-level predictions, based on previous cluster analysis for western Ethiopia (Chapter 1). The advantages of defining homogeneous regions for seasonal prediction at operational (small) scales will be demonstrated by comparing approaches with and without cluster analysis. The combination of objective cluster analysis, (spatially) high-resolution prediction of seasonal precipitation, and a modeling structure spanning statistical and dynamical makes this approach unprecedented and advantageous compared to previous studies.

2.2 Modeling High-resolution Local Seasonal Prediction

To evaluate high-resolution local seasonal precipitation prediction for cluster versus non-cluster cases, statistical models are developed and compared with bias-corrected dynamical model predictions.

Four scenarios are evaluated based on two criteria – (1) *clustered* vs. *non-clustered* and (2) *direct* vs. *indirect*. In the *clustered* case, predictions are produced for each homogeneous region (cluster) given a unique set of predictors. In the *non-clustered* case, the entire study region is considered as one cluster and thus only one set of predictors is utilized for predictions. For the *direct* case, precipitation is predicted directly at local level (grid scale); for the *indirect* case, the average precipitation within each homogeneous region is predicted first (as an intermediary), and then regressed to local-level (grid scale) predictions. Combination of the two criteria forms four scenarios – *clustered direct* (C-D), *non-clustered direct* (NC-D), *clustered indirect* (C-I), and *non-clustered indirect* (NC-I) predictions.

2.2.1 Cluster Analysis

Using a k-means clustering technique, western Ethiopia – the major agricultural region of the country – is clustered into eight homogeneous regions (Figure 1-13), conditioned on the interannual variability of total precipitation in JJAS, the same variable that is to be predicted. Precipitation is based on a $0.1^\circ \times 0.1^\circ$ gridded precipitation dataset from NMA (Dinku et al., 2014), consisting of 7320 grid-cells across 1983 – 2011 (29 years). Given the high-resolution gridded dataset, k-means clustering is performed for the optimal number of clusters. It does not involve any subjective delineation or manual drawing of boundaries between clustered stations or grid-cells; instead, an automatic and objective delineation is performed. Readers are referred to Chapter 1 for more details.

2.2.2 Statistical Modeling Approach

Multiple linear regression (MLR) is favored by many as a statistical modeling approach given its well-developed theory, simple model structure, efficient processing, and yet skillful outcomes (e.g. Omondi et al., 2013; Camberlin and Philippon, 2002; Diro et al., 2008). As mentioned, only a few studies have focused on seasonal precipitation prediction in Ethiopia (Gissila et al., 2004; Block and Rajagopalan, 2007; Korecha and Barnston, 2007; Diro et al., 2008; Diro et al., 2011b; Segele et al., 2015), and almost all of them include the applications of MLR. This chapter also applies MLR to predict seasonal precipitation, yet differentiates from other studies by applying predictions to pre-defined homogeneous regions and further translating to local-level predictions.

Large-scale climate variables are often evaluated as potential predictors in statistical seasonal precipitation prediction models, commonly including sea surface temperatures (SST) in the equatorial Pacific Ocean representing the well-known of the El Nino Southern Oscillation (ENSO) (Stone et al., 1996). For Ethiopia, the ENSO phenomenon is considered a significant indicator of precipitation variability, particularly in the main JJAS rainy season (e.g. NMSA, 1996; Camberlin, 1997; Bekele, 1997; Segele and Lamb, 2005; Diro et al., 2011a; Elagib and Elhag, 2011). In addition to ENSO, the effect of Indian Ocean SST and regional atmospheric pressure systems such as the St. Helena, Azores, and Mascarene Highs also have notable influence on Ethiopia's precipitation variability (e.g. Kassahun, 1987; Tadesse, 1994; NMSA, 1996; Shanko and Camberlin, 1998; Goddard and Graham, 1999; Latif et al., 1999; Black et al., 2003; Segele and Lamb, 2005). Consequently, season-ahead (March-May) or month-ahead (May) large-scale climate variables that are physically relevant in potentially modulating moisture transport to the basin (or cluster) are selected as potential predictors. Four climate variables are selected here for further evaluation based on outcomes of the aforementioned prediction studies: SST, sea level pressure (SLP), geopotential height (GH) at 500mb, and surface air temperature (SAT). All climate variables are from the National Centers for Environmental Prediction and National Center for Atmospheric Research (NCEP/NCAR) reanalysis dataset (Kalnay et al., 1996) at a $2.5^{\circ} \times 2.5^{\circ}$ grid scale.

Predictor selection and statistical modeling are developed according to the following five steps – for the region as a whole (non-clustered) and for each pre-defined cluster (Figure 2-1):

- (1) Precipitation observations for JJAS averaged across the region and each cluster independently are spatially correlated with each global climate variable (e.g. Figure 1-15).

(2) For each spatial correlation, regions with justifiable climatic associations and statistically significant correlations at the 95% level are identified and selected (Table 1-4).

(3) For each climate variable region selected (Table 1-4), data within the region are spatially averaged for 1983-2011 (defined as “pre-predictors”).

(4) Pre-predictors are combined and filtered (for the region or each cluster separately) through principal component analysis (PCA; Jolliffe, 2002).

(5a) The top principal components (PCs) from the PCA are used as predictors – the direct inputs into the principal component regression (PCR) statistical model. For the *direct* case, PCR is used to directly predict the grid-level precipitation; for the *indirect* case, PCR is used to predict an intermediate cluster-level precipitation.

(5b) For the *indirect* case only, cluster-level predictions are regressed to grid-level using MLR.

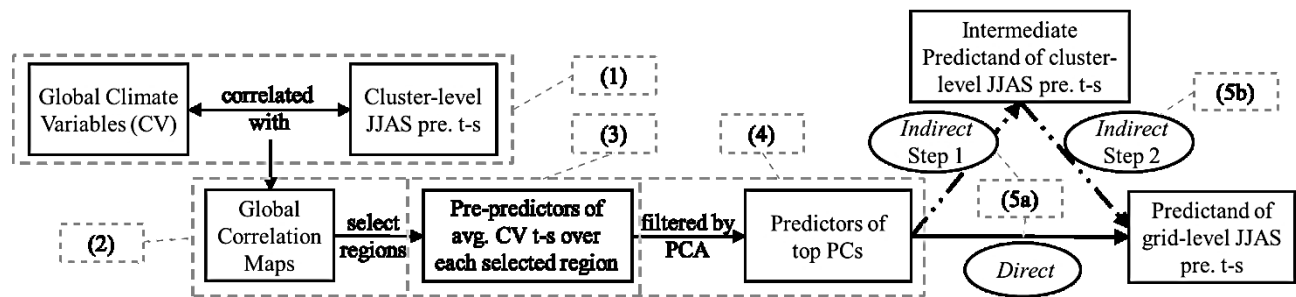


Figure 2-1: Flow chat of data processing for predictors into the statistical model. Numbers framed by dash lines correspond to the procedures listed in the context. *Note: pre. – precipitation, t-s – time series, avg. – average.*

PCA is a common approach in climate modeling to reduce the dimensionality of predictors and remove multi-collinearity, while simultaneously extracting the most dominant signals from the potential predictors, typically reflected in the first few PCs. Since PCA is independent of the

predictand, retaining the first few PCs as predictors, in lieu of the original variables, also helps to reduce artificial prediction skill. A scree test (Jolliffe, 2002) is performed to determine the optimal number of PCs to retain as predictors and the amount of variance explained in the predictors.

PCR is performed in a “drop-one-year” cross-validation mode to reduce over-fitting effects and therefore avoid overestimation of the prediction skills. This requires reconstructing the principal components for the dropped year, and then multiplying the coefficient estimates with each reconstructed PC respectively in order to obtain the final predicted value for the dropped year (e.g. Block and Rajagopalan, 2009; Wilks, 2011). Q-Q plots are evaluated to verify normally distributed residuals (results not included).

For the four scenarios, the models structure are quite similar but have subtle differences which could lead to significantly different outcomes (Table 2-1). Under NC-D (Equation 2-1a, b) and C-D scenario (Equation 2-2a, b), the time series of JJAS seasonal total precipitation in each grid-cell (i.e. at local level) is used as the direct predictand ($Y_{i,t}$); however, NC-D is different from C-D, as the former uses the same predictors (X_t) across all the grid-cells, while the latter uses different predictors according to the cluster to which the grid-cell is assigned ($X_{j,t}$). In the *indirect* case, the cluster-level time series of JJAS seasonal total precipitation (the time series averaged over all grid-cells that belong to a given cluster, $Y_{m,t}$ or $Y_{j,t}$) is first predicted (Equation 2-3a, b and 2-4a, b). The predicted intermediate product ($\tilde{Y}_{m,t}$ or $\tilde{Y}_{j,t}$) is then used as the only regressor in the 2nd step to estimate the grid-level precipitation ($\tilde{Y}_{i,t}$ or $\tilde{Y}_{j,t}$ for every j ; Equation 2-3c, d and 2-4c, d). Again, under C-I the predictors used in the 1st step are different for grid-

cells in different clusters ($X_{j,t}$), while they are the same across all the grid-cells (X_t) under NC-I. Consequently, under C-I the predicted intermediate product includes eight different cluster-level time series ($\tilde{Y}_{j,t}$) associated with grid-cells in different clusters; in contrast, only one cluster-level time series ($\tilde{Y}_t$) consistent across all the grid-cells is produced for NC-I.

Table 2-1: Equations of linear regression panel models under four scenarios

	Non-clustered	Clustered
Direct	$Y_{i,t} = \tilde{c}$ (Equation 2-1a)	$Y_{i \in j,t} = \tilde{c}$ (Equation 2-2a)
	$\tilde{Y}_{i,t} = \tilde{c}$ (Equation 2-1b)	$\tilde{Y}_{i \in j,t} = \tilde{c}$ (Equation 2-2b)
Indirect	$Y_{m,t} = \tilde{c}$ (Equation 2-3a)	$Y_{j,t} = \tilde{c}$ (Equation 2-4a)
	$\tilde{Y}_{m,t} = \tilde{c}$ (Equation 2-3b)	$\tilde{Y}_{j,t} = \tilde{c}$ (Equation 2-4b)
	$Y_{i,t} = \tilde{c} + \tilde{f}_j$ (Equation 2-3c)	$Y_{i \in j,t} = \tilde{c} + \tilde{f}_j$ (Equation 2-4c)
	$\tilde{Y}_{i,t} = \tilde{c} + \tilde{f}_j + \tilde{f}_t$ (Equation 2-3d)	$\tilde{Y}_{i \in j,t} = \tilde{c} + \tilde{f}_j + \tilde{f}_t$ (Equation 2-4d)

where Y - predictand of JJAS seasonal total precipitation; X - PC predictors; ε, ν - error terms; $\tilde{Y}$ - predicted values of JJAS seasonal total precipitation; $\tilde{c}, \tilde{f}_j, \tilde{f}_t, \tilde{f}_{jt}$ - estimated coefficients; i - grid-cell index; t - time (year) index; j - cluster index; $i \in j$ - grid-cell i that belongs to cluster j ; m - mean over entire study region that is equivalent to having only one cluster.

2.2.3 Dynamical Modeling Approach

The North American Multi-Model Ensemble (NMME; Kirtman et al., 2014) is an experimental multi-model system consisting of coupled dynamical models from various modeling centers in North America that includes seasonal predictions. The hindcasts of monthly mean precipitation are accessible through the International Research Institute for Climate and Society (IRI) website (<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/>). To compare with statistical model predictions, NMME JJAS seasonal precipitation predictions ($1^\circ \times 1^\circ$ grid-cells) are extracted

from model ensembles that cover the same time period (1983 – 2011), same geographic region (western Ethiopia), and same lead time (predictions made on June 1). A subset of 10 NMME models meet these criteria and are retained for further evaluation: (1) COLA-RSMAS-CCSM3, (2) COLA-RSMAS-CCSM4, (3) GFDL-CM2p1, (4) GFDL-CM2p1-are04, (5) GFDL-CM2p5-FLOR-A06, (6) GFDL-CM2p5-FLOR-B01, (7) IRI-ECHAM-AnomalyCoupled, (8) IRI-ECHAM-DirectCoupled, (9) NASA-GMAO, (10) NCEP-CFSv2.

The NMME predictions for each of the 10 models are bias-corrected through probability mapping (e.g. Block et al., 2009; Teutschbein and Seibert, 2012; Chen et al., 2013), subject to the observational dataset from NMA. This is performed on a grid-cell by grid-cell basis on standardized data (the NMME dataset is reshaped to $0.1^\circ \times 0.1^\circ$ grid-cells to match the observational NMA dataset grid-cell size). The basic steps include:

- (1) Fit gamma distributions to each observed and NMME grid-cell respectively; for NMME this is by model using all ensemble members available. (Goodness-of-fit tests indicate gamma distributions fit well; results not shown).

- (2) Translate into cumulative distribution functions (CDF).

- (3) For any given probability, the corrected prediction value can be found by mapping from the NMME CDF to the observed CDF and applying the inverse gamma distribution. Repeat for each grid-cell at $0.1^\circ \times 0.1^\circ$ resolution.

As a result, the gamma CDF of predictions and observations approximately match after correction (Figure 2-2a). Additionally, each ensemble still retains its variability over time, though the overall ensemble mean is shifted to closely match observation (Figure 2-2b).

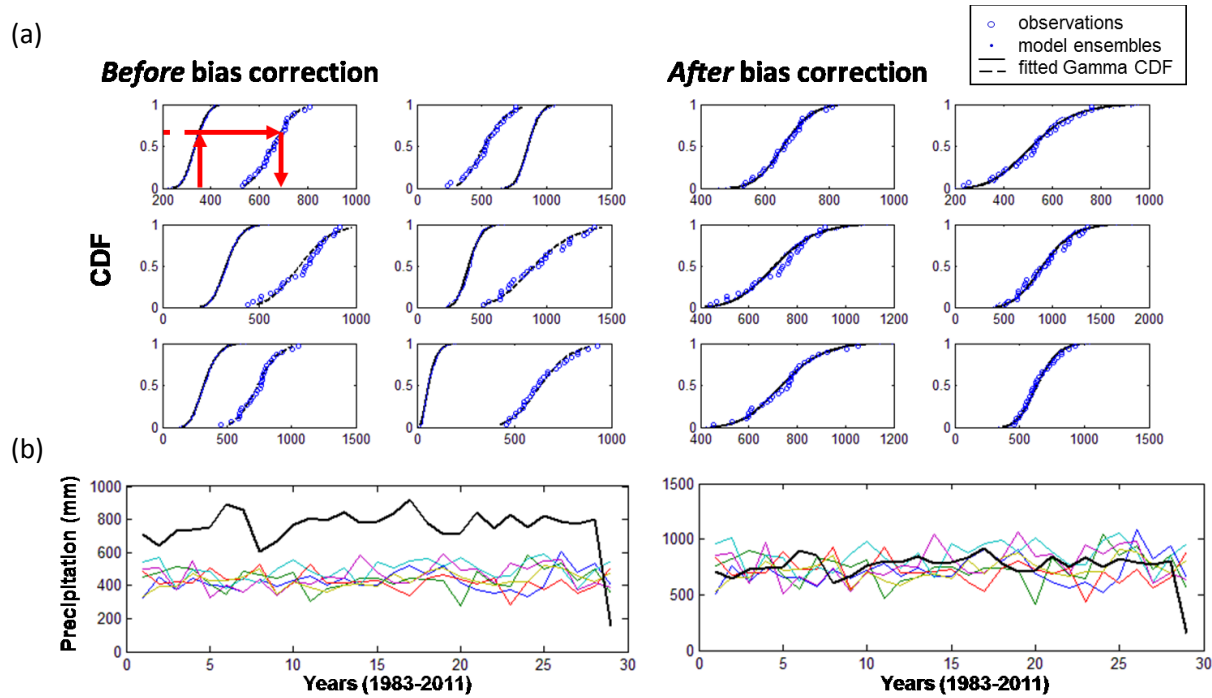


Figure 2-2: (a) bias correction of NMME predictions using probability mapping; (b) precipitation time series from NMME (colored lines) before and after correction, compared to observations (black line). Examples are shown for randomly selected six grid-cells.

2.2.4 Performance Metrics

Pearson correlations are used to measure the standardized covariance between observations and predictions. Ranked probability skill scores (RPSS; Wilks, 2011) are also evaluated to determine categorical skill based on probabilistic predictions. Here, the data are split into three equal terciles representing below-normal, near-normal, and above-normal conditions. A perfect prediction yields an RPSS of 100%, and a prediction with less skill than climatology (long-term averages) yields an RPSS of less than zero. Median RPSS values from all 29 years are reported.

Overall model superiority is evaluated by Akaike information criterion (AIC), Bayesian information criterion (BIC), and generalized cross validation (GCV) scores (Craven and Wahba, 1979; Manning et al., 2008b). All metrics reward model parsimony by penalizing models with a

larger number of predictors. Smaller AIC, BIC and GCV scores are preferred. The equations used to calculate AIC, BIC, and GCV, respectively, are given by:

$$AIC = N \times \log\left(\frac{RSS}{N}\right) + 2 \times K \dots\dots\dots \text{(Equation 2-5)}$$

$$BIC = N \times \log\left(\frac{RSS}{N}\right) + \log(N) \times K \dots\dots\dots \text{(Equation 2-6)}$$

$$GCV = \frac{RSS}{N \times \left(1 - \frac{K}{N}\right)^2} \dots\dots\dots \text{(Equation 2-7)}$$

where N is the number of years (29 years), K is the number of predictors used in the regression, and RSS is the residual sum of squares (equal to the difference between observations and predictions in each year squared, summed over all the years).

2.3 Results

2.3.1 Statistical Model Predictions

Top two PCs are retained as the predictors for each cluster in the *clustered* case and the top three PCs are retained in the *non-clustered* case, with an average of approximately 70% total variance explained.

For the C-I and NC-I scenarios, cluster-level precipitation is first predicted, which is subsequently used as a regressor to predict local-level (i.e. grid-level) precipitation. The cluster-level observations are captured closely by the modeled precipitation over time, with most of the observations falling within the predicted 95% confidence envelope (Figure 2-3). Strong positive

correlations with observations – ranging from 0.68 to 0.84 over 29 years – are shown across all clusters in the *clustered* case and the *non-clustered* study region (Table 2-2). Additionally, all RPSS values are positive, indicating superior prediction skill over climatology. Among all clusters, *Cluster 5* in agriculturally rich central-northwestern Ethiopia (Figure 1-13) performs best, with correlation and RPSS values of 0.84 and 74.3%, respectively.

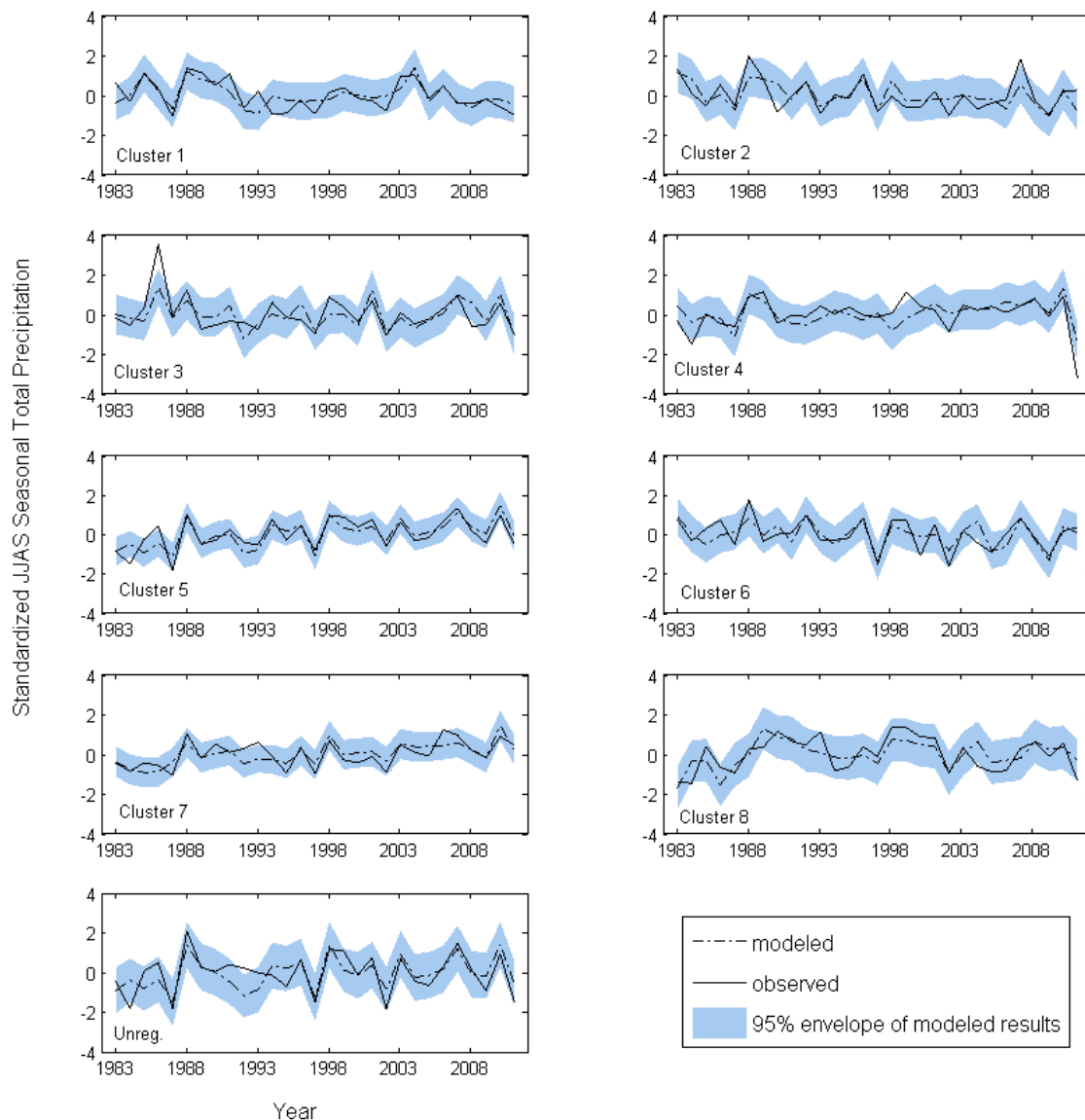


Figure 2-3: cluster-level predictions and observations under C-I and NC-I scenario, with drop-one-year cross-validation. The 95% envelope shows the 95% confidence interval constructed using model errors (extended from Figure 1-17).

Table 2-2: Correlation coefficients (Corr.) and RPSS for predictions (drop-one-year cross-validated) at cluster level compared to observations under C-I and NC-I scenario (extended from Table 1-5).

Cluster	C1	C2	C3	C4	C5	C6	C7	C8	Non-cluster
Corr.	0.741	0.695	0.711	0.683	0.838	0.744	0.751	0.699	0.739
RPSS (%)	45.23	26.04	36.16	19.82	74.30	5.44	51.91	48.21	48.72

At the grid-scale correlations between predictions and observations are clearly superior for the *clustered* case versus the *non-clustered* case (Figure 2-4). Some parts of the region reach a correlation of 0.9, such as central-northwestern Ethiopia, which is consistent with the region of high cluster-level prediction skill (*Cluster 5*). The average correlation over all the grid-cells is approximately 0.51 (*direct*) and 0.53 (*indirect*) for *clustered* predictions, compared to 0.24 (*direct*) and 0.27 (*indirect*) for the *non-clustered* predictions (Table 2-3), although spatial differences are clearly apparent (Figure 2-4). In addition to higher average correlations, standard deviations of correlations are lower in the *clustered* case than in the *non-clustered* case, indicating a more concentrated correlation distribution at higher values. The percent of grid-cells with positive correlations passing the 95% significance test increases from approximately 30% in the *non-clustered* case to more than 80% in the *clustered* case (Table 2-3).

Similarly findings are evident by evaluating the RPSS. The *non-clustered* predictions are modestly skillful, particularly for the same region of central-northwestern Ethiopia (Figure 2-5), with an average RPSS of approximately 5.4% (both *direct* and *indirect*) over entire study region, however RPSS values improve nearly fourfold to 19.2% and 26.5% in the *clustered* case (*indirect* and *direct*, respectively). Additionally, the percentage of grid-cells with positive RPSS values, indicating skillful predictions, reaches 84.4% - 90.0% in the *clustered* case (Table 2-3).

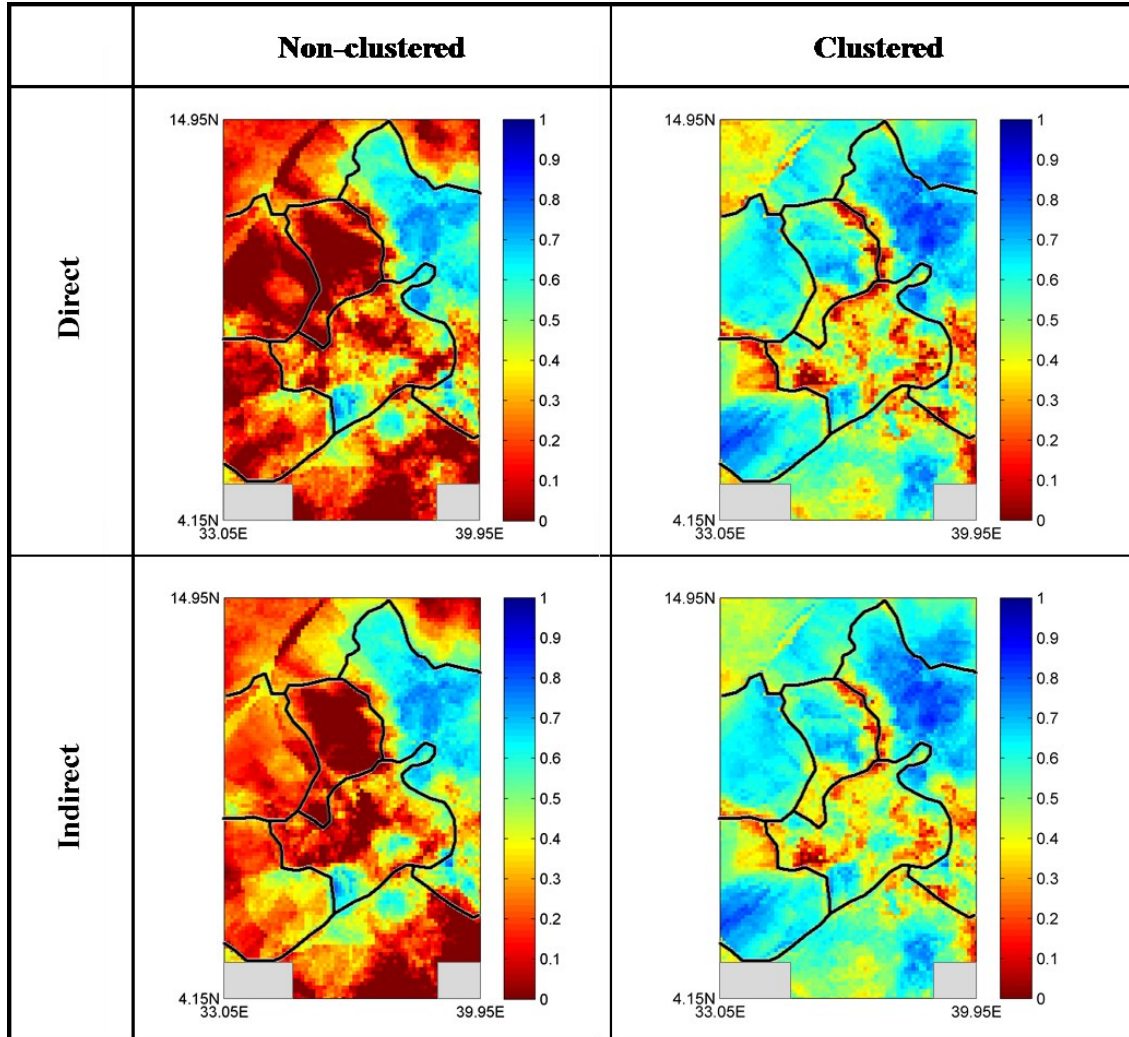


Figure 2-4: Pearson correlations between grid-level observations and predictions under four scenarios, with the clustering boundary delineated roughly in black.

At the grid-scale, predictions by the *indirect* approach generally outperform *direct* approach predictions, based on AIC, BIC and GCV values (Table 2-4), as well as correlation and RPSS values (Table 2-3). Using the predicted cluster-level precipitation to predict grid-level precipitation (the *indirect* case) appears to help reduce the effect of over-fitting and smooth grid-scale noise. From another perspective, the results also suggest that precipitation signals at the regional scale are better explained by large-scale climate variables, while at highly localized

scales the signal is less evident. Obviously, however, this is dependent on cluster size and the degree of spatial coherence within each cluster, as demonstrated in this chapter.

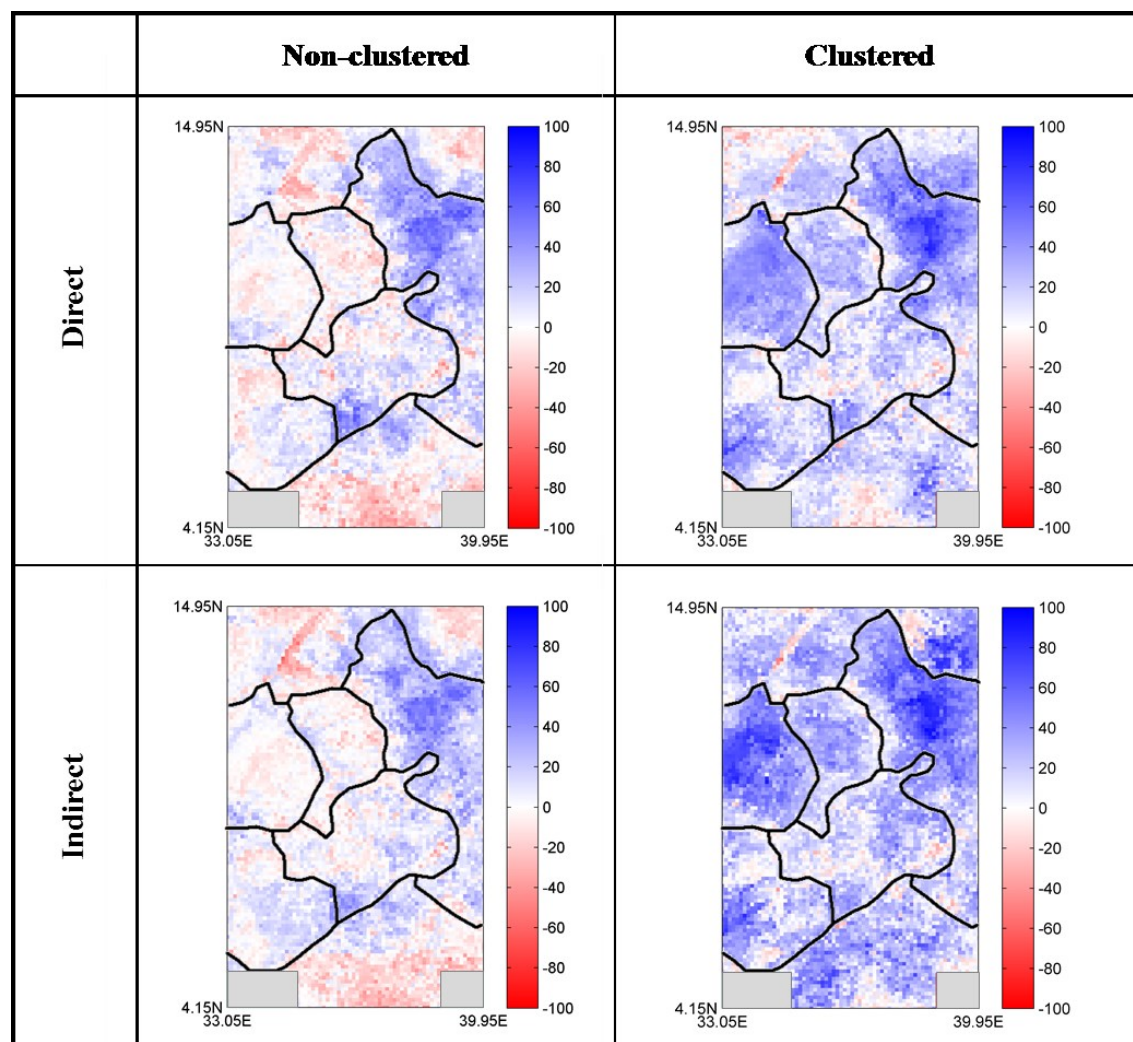


Figure 2-5: grid-level RPSS (%) under four scenarios using climate variables as predictors, with the clustering boundary delineated roughly in black.

Table 2-3: Grid-level Pearson correlation and RPSS statistics

Statistical Model	Grid-level correlations			Grid-level RPSS		
	mean	stdev	significant corr %	mean (%)	stdev (%)	positive RPSS %
NC-D	0.237	0.245	28.1%	5.42	18.46	58.8%
NC-I	0.272	0.247	32.3%	5.32	17.09	60.7%
C-D	0.509	0.172	80.8%	19.16	19.18	84.4%
C-I	0.532	0.146	87.1%	26.47	21.47	90.0%
Dynamical Model						
(9) NASA-GMAO	0.300	0.149	36.1%	2.32	21.20	54.3%
(10) NCEP CFSv2	0.310	0.155	37.3%	3.66	16.61	61.0%

Table 2-4: grid-level AIC, BIC, and GCV value statistics

Statistical Model	AIC		BIC		GCV	
	mean	stdev	mean	stdev	mean	stdev
NC-D	281.05	17.86	285.15	17.86	2.00E+04	1.46E+04
NC-I	277.47	17.37	280.20	17.37	1.73E+04	1.20E+04
C-D	272.79	15.38	276.89	15.38	1.41E+04	7.53E+03
C-I	270.05	15.33	272.78	15.33	1.27E+04	6.84E+03

2.3.2 Dynamical Model Predictions

The RPSS values based on the prediction ensembles of each dynamical model improve significantly after bias correction, however, the median RPSS values over all the grid-cells are still close to zero (Figure 2-6). The last two models, (9) NASA-GMAO and (10) NCEP-CFSv2, are the only two models showing a positive mean RPSS value of 2.32% and 3.66%, respectively (Table 2-3). All other dynamical models have a negative grid-level RPSS mean ranging from -26.2% to -1.78% (table not shown). The dynamical models (9) and (10) also exhibit generally higher grid-level correlations over the study region (Figure 2-7), with an average value of 0.30 and 0.31, respectively (Table 2-3). The overall performance of the two best dynamical models,

however, is still inferior to that of the *clustered* statistical models, in terms of the prediction skills evaluated by the correlation and RPSS metrics (Table 2-3).

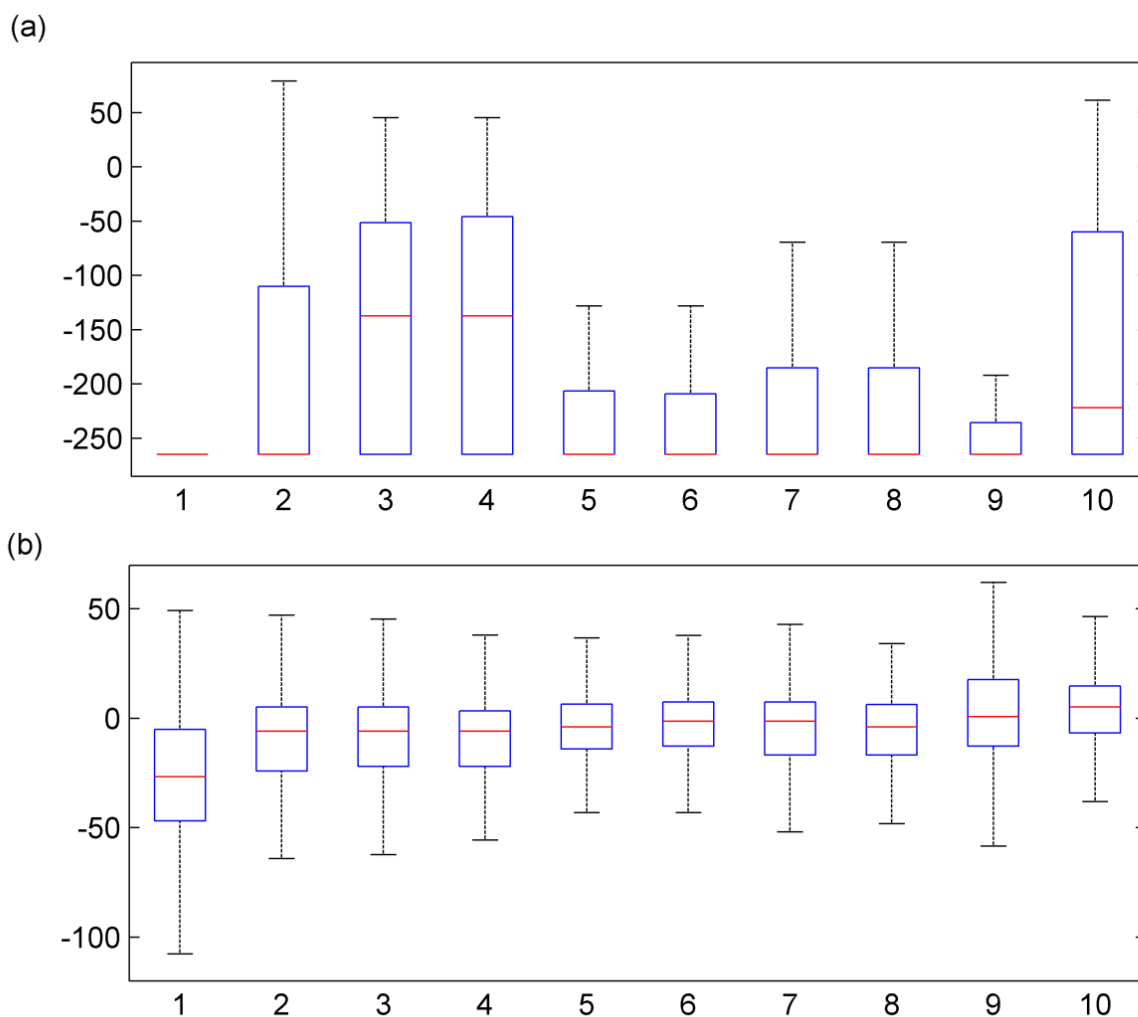


Figure 2-6: Boxplots of grid-level RPSS (%) for 10 dynamical models from NMME (a) before and (b) after bias correction, labeled with the same number as listed in the context. *Note: For each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, and the whiskers extend to the most extreme data points not considered outliers (outliers not shown).*

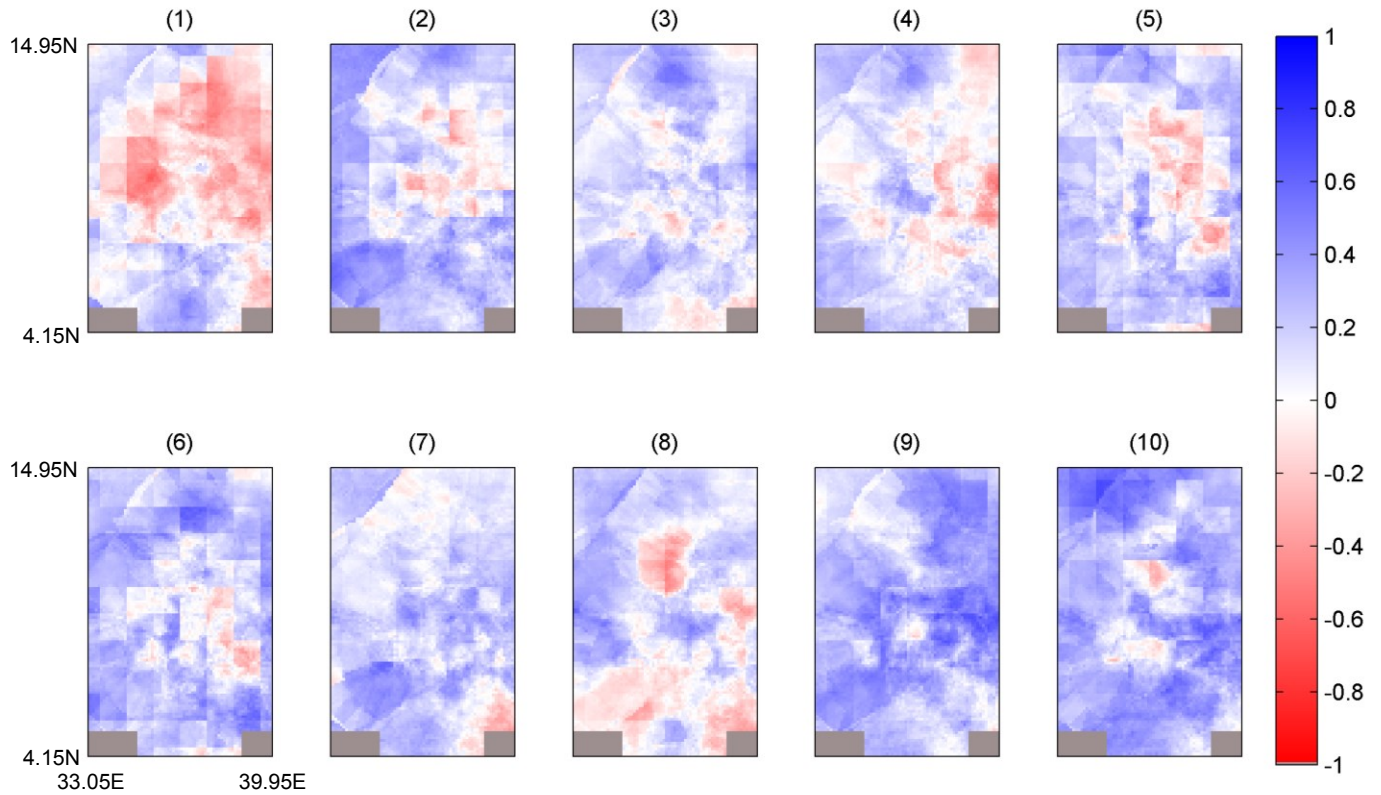


Figure 2-7: Pearson correlations between grid-level observations and ensemble mean of bias-corrected predictions for 10 dynamical models from NMME, labeled with the same number as listed in the context. *Note that the scale ranges from -1 to 1.*

2.4 Conclusions and Discussion

This chapter illustrates that using prior season large-scale climate information can lead to skillful and credible statistical precipitation predictions at local scales under a *clustered indirect* approach in western Ethiopia. To our knowledge, this is the first study predicting seasonal precipitation at high resolution in western Ethiopia. At the regional scale, the approach demonstrated in this chapter shows promise, particularly compared to current NMA operational forecasts, which have a weak positive skill relative to climatology (Korecha and Sorteberg, 2013). The *spatially high-resolution* predictions here are also advantageous to *spatial average* predictions in previous studies (Gissila et al., 2004; Block and Rajagopalan, 2007; Korecha and

Barnston, 2007; Diro et al., 2011b; Segele et al., 2015), given the preliminary cluster analysis process. An exclusive set of predictors is used for each cluster, which contributes to advanced prediction skills at cluster and grid level, compared to *non-clustered* prediction. Grid-level prediction under the *indirect* case reduces the overfitting effect, which is captured by cross-validation, relative to the *direct* case. This improvement certainly requires an appropriate delineation of homogeneous rainfall regions.

Although predictions from the statistical model *clustered* approach are superior to all dynamical predictions in this chapter, improvements in dynamical models continues, and their application to seasonal precipitation prediction is likely to grow (e.g. Palmer et al., 2004; Saha et al., 2006; Lim et al., 2009). Multi-model combinations of statistical and dynamical models were also investigated for potential improvement of prediction skills through pooling (assign equal weights to all model candidates), linear regression and Bayesian Model Averaging (BMA; Raftery et al., 1997). Three candidate models, namely the best statistical model (C-I) and two best dynamical models (NASA-GMAO and NCEP-CFSv2) are selected; however, the overall performance was noticeably lower compared to the single statistical model C-I (results not shown).

Although *clustered* statistical model predictions are promising overall, it is worth noting that relatively poor prediction performance is evident in some locations. One such place is along cluster boundaries, where assignment of grid-cells to one certain cluster versus the neighboring cluster is almost arbitrary, and clearly less certain than grid-cells falling within the central parts of clusters (Figure 2-4 and 2-5). Poor prediction skill is also evident in some of the mountainous regions of the study area, where the hydroclimatic processes that produce precipitation are likely driven by orographic and other factors rather than large-scale climate variables. To test the

prospects for improving prediction performance, season-ahead local variables such as soil moisture, spring rains, etc. were investigated, however, no significant improvement was found for the *clustered* case, and the correlations actually deteriorate for the *direct* case. Adding local predictors in this case may introduce more noise and cause over-fitting. It is also possible to explore a different number of clusters, compare prediction skills, and combine the prediction if necessary for improving grid-level predictions. While this is non-trivial, involving both cluster analysis and prediction, and beyond the scope of the study – prediction *given* homogeneous regions, the analysis is not performed here.

Additional prediction features also warrant future attention, including longer prediction lead times and evaluation of other characteristics (e.g. intra-seasonal dry spells, seasonal onset or cessation, etc). Improving predictive capabilities may not be a complete panacea, but it can continue to be an important part of decisions-maker's portfolios as they cope with hydroclimatic variability.

Chapter 3 Integrating Predictive Information into an Agro-economic Model to Guide Agricultural Planning

3.1 Introduction

A number of studies explore the expected economic value of seasonal climate forecasts for different agricultural systems at different scales; a summary of relevant literature can be found in (Meza et al., 2008). Most of the studies focus on the value of El Niño Southern Oscillation (ENSO) forecasts (e.g. Hammer et al., 1996; Letson et al., 2005; Marshall et al., 1996; Messina et al., 1999; Solow et al., 1998; Adams et al., 2003), while only a few investigate the value of seasonal precipitation predictions and none address country level outcomes (Wilks and Murphy, 1986; Katz et al., 1987; Mjelde et al., 1988; Jones et al., 2000). Although the majority of studies focus on crops under rainfed conditions, they all emphasize intensification of modern technology (e.g. fertilizer, pesticide, degree of mechanization etc.) and commercial agriculture in general; subsistence agriculture is consistently omitted (Meza et al., 2008).

In this chapter, we apply an *ex-ante* country-level evaluation of seasonal precipitation prediction using an equilibrium model to simulate Ethiopia's national economy accounting for interannual climate variability and prediction-guided agricultural responses. It provides perspectives at an aggregate scale (i.e. overall welfare to society) and across various economic indices (e.g. gross domestic product, calorie consumption, poverty rate, etc.). Agricultural production, particularly subsistence farming, the foundation of the country's economy, is explicitly modeled in this chapter. With 80% of its population living in rural areas and engaged in farming, there is a high susceptibility to the impacts of climate variability (Oram, 1989; Dixon and Segerson, 1999;

Hansen, 2002). The motivation of this chapter is to present an innovative means of evaluating predictive information, which can serve as a foundation for communication, decision making, and strategic planning.

3.2 Methods

Ethiopia's economy-wide multi-market model (EMM; Diao and Pratt, 2007), originally developed and calibrated by the International Food Policy Research Institute (IFPRI), is modified to (1) simulate the zonal to country-level economy with a dynamic (varying) climate, and (2) evaluate predictive information (presented in Chapter 2) by comparing various economic outputs for prediction-guided agricultural responses versus baseline agricultural management.

3.2.1 Ethiopia's Economy-wide Multi-market Model (EMM)

The EMM (Diao and Pratt, 2007) is an economy-wide, multi-market model, with detailed structure of agricultural sector and additional two aggregated sectors representing industry and services. Specifically, supply and demand of 32 agricultural commodities, such as crops and livestock, and 2 aggregated non-agricultural commodities are modeled at zonal level. Therefore, each of the 56 administrative zones has its own representative supply function to capture producer's response to market. The supply for crops is modeled with crop yield multiplied by crop area, both of which are functions of price while yield is also a function of climate factors. The demand function is also disaggregated to the zonal level, reflecting consumer's demand for each commodity given its market price and per capita income. Price elasticity including own-price and cross-price elasticities, and income elasticity are varied across zones for each

commodity given different income levels and consumption patterns. Per capita income is endogenously determined by supply revenue divided by population, which allows supply and demand being linked at zonal level. Since the EMM is benefit-only with intermediate inputs and their costs omitted, producer price is therefore adjusted to represent the value added. Consequently, the aggregation of zonal supply at its value-added prices equals the gross domestic product (GDP), which is consist of agricultural GDP and non-agricultural GDP. Other output variables such as poverty rate and calories per capita per day are also calculated given zonal income levels and food consumption.

Multi-markets linkages are modeled through price margins for each zone to the price in the national central market at Addis Ababa. The price margins between markets are determined according to the distance of each zone to Addis Ababa, representing the basic transportation costs. Food surplus zones with supply greater than demand would face a lower price than the price at central market, with the difference being the marketing margins at equilibrium. Food deficit zones, in contrast, would endure a higher price, which equals to the price at central market plus the transportation costs. The model also captures international imports and exports with the assumption that the domestic and international commodities are perfect substitutes but distinguished by transportation and other market costs. For example, if the supply of maize decreases along with an increasing domestic price, the import of maize from other countries is only profitable when the domestic price exceeds the import parity price plus any transaction costs. Similarly for export, the domestic price of one commodity has to be low enough to trigger the exports. With imports and exports, the aggregated supply and demand of each commodity reaches equilibrium at national level.

The model is originally calibrated for the base year 2003, using data from national household surveys, agricultural sample surveys, geographic information systems, and other national and regional data. More detailed information about the EMM can be found in Diao and Pratt (2007).

3.2.2 Climate Yield Factor (CYF) and Incorporating Climate Variability into the EMM

Climate yield factor (CYF; Block et al., 2008), which represents the overall effect of climate on crop yields, is calculated based on procedures provided by the United Nations Food and Agriculture Organization's (FAO) Publication 33, Yield Response to Water (Doorenbos and Kassam, 1979), and Publication 56, Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements (Allen et al., 1998). The CYF values of six staple crops in Ethiopia, including teff, maize, wheat, sorghum, millet, and barley, are explored (barley shares the same CYF values as sorghum given their similarity). Cash crops, such as pulses, oil crops, and enset, are assumed to have the highest CYF values over all six staple crops explored for each year.

To calculate the CYF, monthly climate data including elevation, cloud cover, temperature, diurnal temperature range, vapor pressure, and wind from Climate Research Unit (CRU)'s gridded dataset (Harris et al., 2014) are utilized to calculate reference evapotranspiration (ET_0) using the Penman-Monteith method (Allen et al., 1998). Consequently, each grid-cell has a unique ET_0 value for each month in each year from 1983 – 2011. Potential crop evapotranspiration (ETC) for each crop in each grid-cell can then be obtained using ET_0 multiplied by a crop-specific empirical constant K_s . Gridded monthly precipitation observations from the National Metrology Agency (NMA) of Ethiopia (Dinku et al., 2014) and CRU (Harris et al., 2014) covering all zones, and soil data from the FAO Digital Soil Map of the World

(FAO-UNESCO, 1988) are used to obtain actual evapotranspiration (ETA) through a water balance equation. The CYF is determined according to the ratio of ETA over ETC and a crop's sensitivity to limited water availability. A lower ETA/ETC ratio and higher sensitivity to water stress (K_y) would produce a relatively low CYF value, indicating a greater impact on yield due to water scarcity. The specific equation is provided below:

$$CYF = 1 - K_y \cdot \left(1 - \frac{ETA}{ETC}\right) \dots\dots\dots \text{(Equation 3-1)}$$

where K_y is a crop-specific constant (FAO Publication 33) - a higher value indicates more vulnerable to water scarcity.

The resultant CYF ranges from zero to one, where $CYF = 1$ implies that yields are not limited by water stress, while limitation by other factors such as pests, soil fertility, management skills are still possible. A $CYF = 0.9$ indicates a 90% yield based on water availability. A $CYF = 0$ indicates crop failure. Note that the calculation is performed for each crop stage spanning different months – vegetative, flowering, yield formation, harvest – given different K_s and K_y values. The lowest CYF across all stages in one annual cycle is retained as the final CYF. Eventually, each grid-cell has one CYF value for each crop in each year.

Since the EMM operates at a zonal level, the gridded CYF values are converted to zonal values by overlaying a shape file of all zones on the gridded region and calculating the weighted average of the grid-cells that fall into each zone. The weight depends on the percent of grid-cell area that falls within the zone over the total zonal area. Hence, weights of all the grid-cells that fall into a zone sum to one, and a grid-cell which partially falls inside a zonal boundary has a lower weight than a grid-cell which completely falls inside the zone.

The original EMM was developed using static historical average climate to simulate the zonal to national economy across years. In this chapter, the CYF values in each year are used to incorporate dynamic climate variability into the EMM, thus the EMM is recalibrated using the CYF values specifically in the base year 2003, instead of the historical average. Simulations with a “shock” on climate – different CYF values from the base year – are generated for each of the other years. That is, with all other inputs fixed, it can illustrate the economic influence exclusively due to climate variability. In such cases, each year is an independent experiment without accounting for growth in population, crop area, etc.

3.2.3 Integrating Seasonal Prediction to Guide Agricultural Planning – Altering Area Planted for Two Main Crops in Ethiopia

Seasonal prediction (June to September, or JJAS) of precipitation spanning 29 years (1983-2011) in western Ethiopia developed in Chapter 2 is applied here, using a categorical prediction of above-normal, near-normal, and below-normal intervals to represent wet, normal, and dry conditions. Intervals are selected such that 10 of the observed 29 years fall into above-normal and below-normal intervals respectively, while the remaining 9 years fall into the near-normal interval. This categorical prediction is consistent with the current operational seasonal forecasts issued by NMA in Ethiopia (Korecha and Sorteberg, 2013).

Maize and teff are the most common staple crops in Ethiopia and distinct from each other in terms of their sensitivity to water stress – maize is more vulnerable to drought, while teff can better survive under drought conditions. However, maize has a higher economic value per unit of planting area than teff under ideal climate conditions. Thus, maize may be more desirable in wet

years (above-normal conditions), while teff may be more valuable in dry years (below-normal conditions) when maize crop failure is likely to occur. To test the sensitivity of reallocating a certain percentage of maize to teff or vice-versa, a range of percentages from 5%, 10%, ..., to 95% for “maize-to-teff” and “teff-to-maize” scenarios are explored. The total area planted, according to the baseline, remains unchanged. Note that the fixed total area refers to the area without modern technology inputs, such as irrigation, fertilizer, pesticide, and improve seed, where precipitation is the major factor affecting the yield. If benefits in excess of the baseline accrue under the scenarios tested, a skillful seasonal prediction may then help to prescribe the proper shift in crop allocation. Note that all scenarios apply to the country as a whole, that is, shifts in crop allocation are applied across all zones equivalently.

3.2.4 Overall Methodology Structure

In summary, a CYF is calculated for each staple crops to reflect the overall influence of climate variables (mainly precipitation) on crop yield. Since the EMM is based on administrative zonal levels, grid-level CYF values given by gridded climate and soil data are converted to zonal values. Baseline outputs are generated given “business as usual” with dynamic CYF values over years. Predictions, in contrast, allow farmers the opportunities to take alternative actions, aiming to gain anticipated benefits or reduce potential losses. By prescribing farmers’ action to allocate farmland to different types of crops according to the predicted climate condition, the EMM can generate the net benefits given predictive information and compare it to “business as usual” (Figure 3-1).

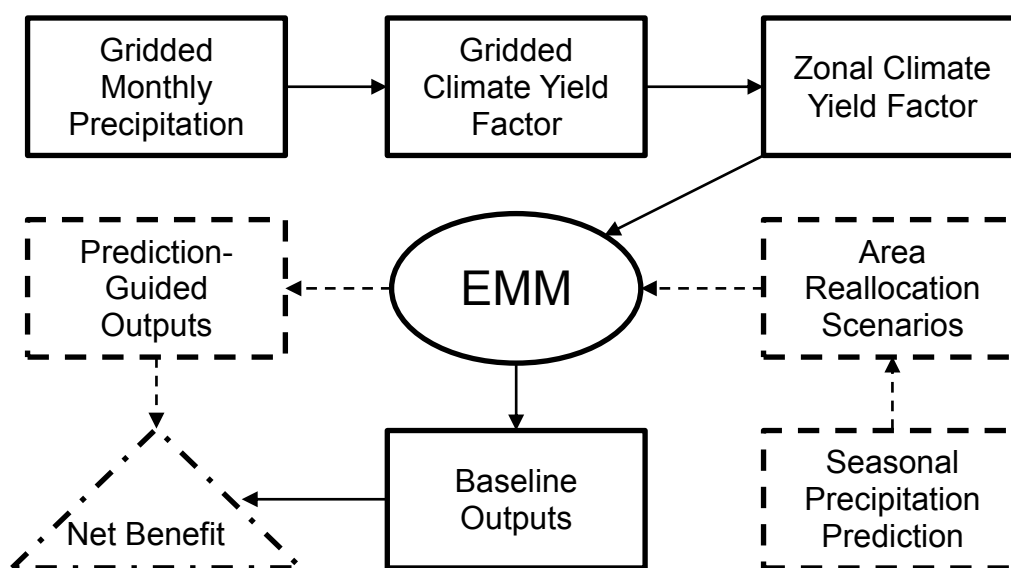


Figure 3-1: Flow chat of methodology. Solid line arrows indicate the process of baseline simulation and dashed line arrows indicate the add-on prediction-guided simulations.

3.3 Results

3.3.1 CYF and Crop Land Devoted to Maize and Teff

The spatial average zonal CYF values for six staple crops show that maize is most sensitive to water stress and it also has the largest variability over years for different climate conditions (Figure 3-2). The lowest spatial average CYF for maize occurs in 2002, a notorious drought year in Ethiopia. The value falls below 0.5, indicating the yield of maize in 2002 would be less than 50% of the yield with sufficient water inputs. In relative wet years, the spatial average CYFs for maize range from 0.65 to 0.7; however, spatial variability of zonal CYF values exists with some wet zones having a CYF value of maize close to one, while zero for other typically dry zones (results not shown). This also applies to other crops but less drastically. In contrast to maize, teff is less sensitive to drought with the highest CYF values among the six staple crops in almost every year (values range from around 0.74 to 0.87; Figure 3-2).

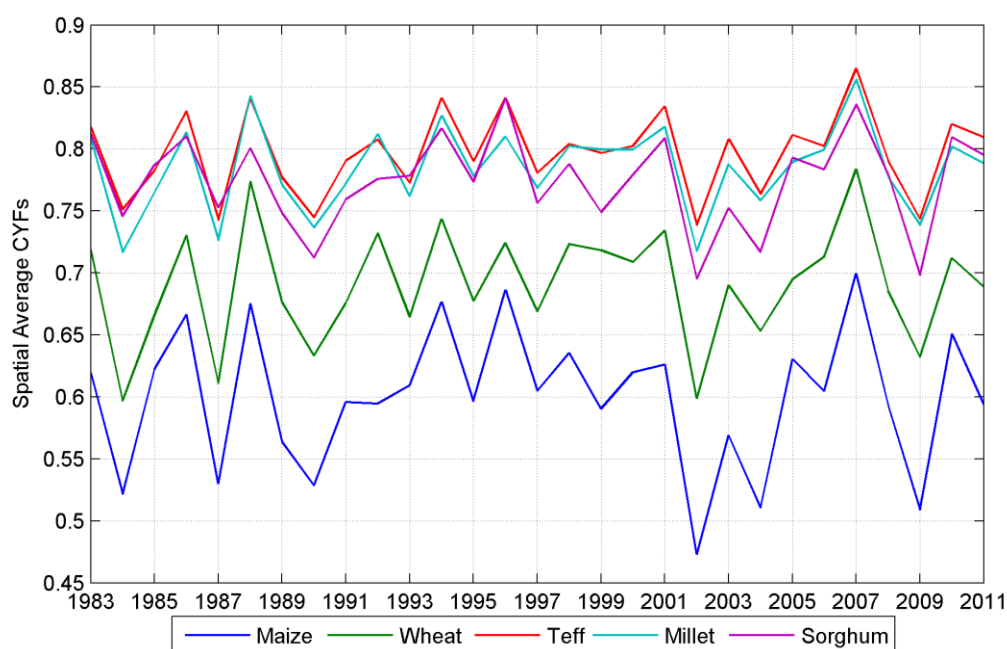


Figure 3-2: Spatial average zonal CYF values over all 56 zones for six staple crops, with barley assumed to share the same CYF values as sorghum.

As mentioned, crop area allocations in Ethiopia are based on calibration results for 2003 (Figure 3-3). Among all the crops, teff is the most common type, occupying 20% of the total crop area, followed by maize at 16%. For teff, 54% of its planting area does not utilize any modern technological inputs – irrigation, fertilizer, pesticide, improved seed, or any combination of them, while for maize a smaller percentage (43%) uses no modern technological inputs (Figure 3-3). Therefore, approximately 18% of the total crop area in Ethiopia is available for reallocating between teff and maize under the proposed scenarios.

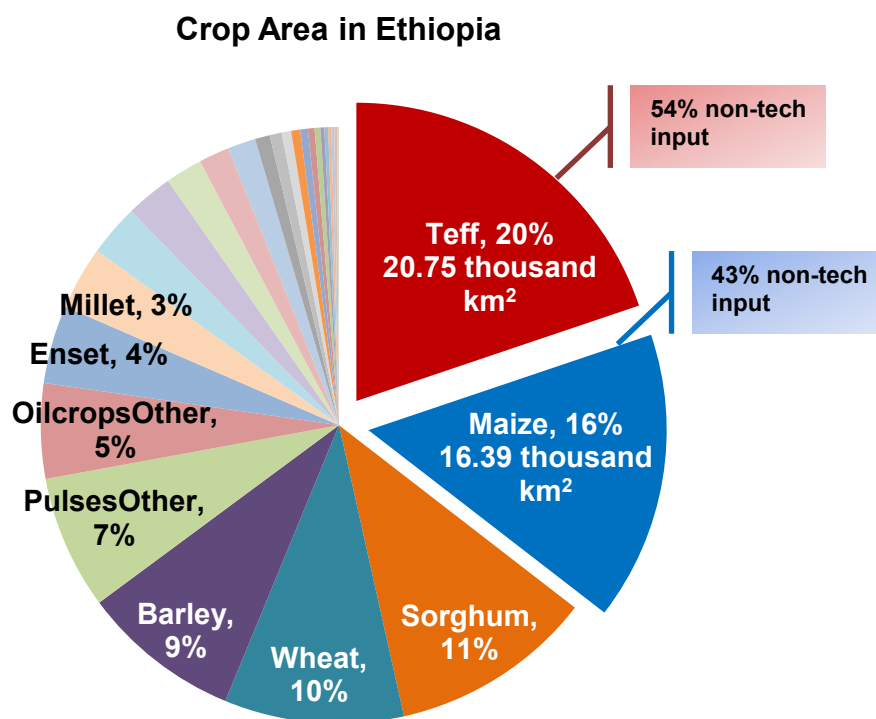


Figure 3-3: Crop area allocation in Ethiopia. “PulsesOther” represents pulses except beans and peas. “OilcropsOther” represents oil crops except sesame, groundnut, and rape seed.

3.3.2 Incorporating Climate Variability – Baseline Simulation

The GDP is strongly positively correlated with JJAS total precipitation ($corr. = 0.679$, $p\text{-value} < 0.0001$; Table 3-1). Note that all GDP in the chapter refers to *real* GDP, which is based on prices *in the base year (2003)*. The agricultural GDP, which is a large portion of total GDP, has an even higher correlation with precipitation than total GDP, as the former is more directly affected by climate conditions. Similarly, grain GDP which is based on mostly staple crops has a slightly higher correlation than agricultural GDP, given an even stronger connection with climate (Table 3-1). Unsurprisingly, in a wet year with higher JJAS precipitation, the total supply of commodities is higher due to preferable climate conditions, as is the GDP. Across the years explored, total GDP can vary by as much as 0.72 billion USD, fully attributable to precipitation

reduction, *ceteris paribus* (i.e. with all other parameters held constant) (e.g. comparing year 1996 to 2002; Figure 3-4).

Table 3-1: Pearson correlation coefficients between JJAS total precipitation and different economic indicators over years, respectively.

	GDP	Agricultural GDP	Grain GDP	Calorie per capita per day	Poor Population	Price of Maize	Price of Teff	Import of Wheat
Corr. with precipitation	0.679***	0.684***	0.686***	0.637**	-0.728***	-0.563*	-0.228	-0.690***

* $p < .002$, ** $p < .0005$, and *** $p < .0001$ under two-tailed t test.

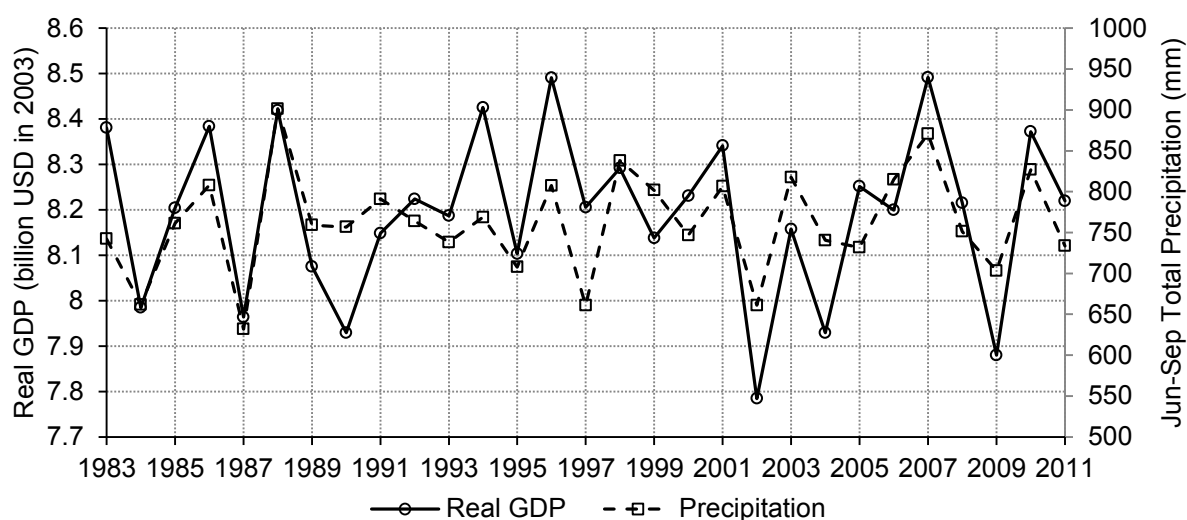


Figure 3-4: National real GDP and spatial average JJAS total precipitation over simulated years.

Total calories per capita per day also show a strong and positive correlation with JJAS total precipitation (Table 3-1), indicating that normal and wet climate conditions brings more food and food consumption. This applies to both food deficit and surplus zones in general, though some spatial variability exists. The deficit zones with food supply less than demand face a lower level of calorie consumption than surplus zones, with a mean difference of 212 Cal per capita per day (Figure 3-5).

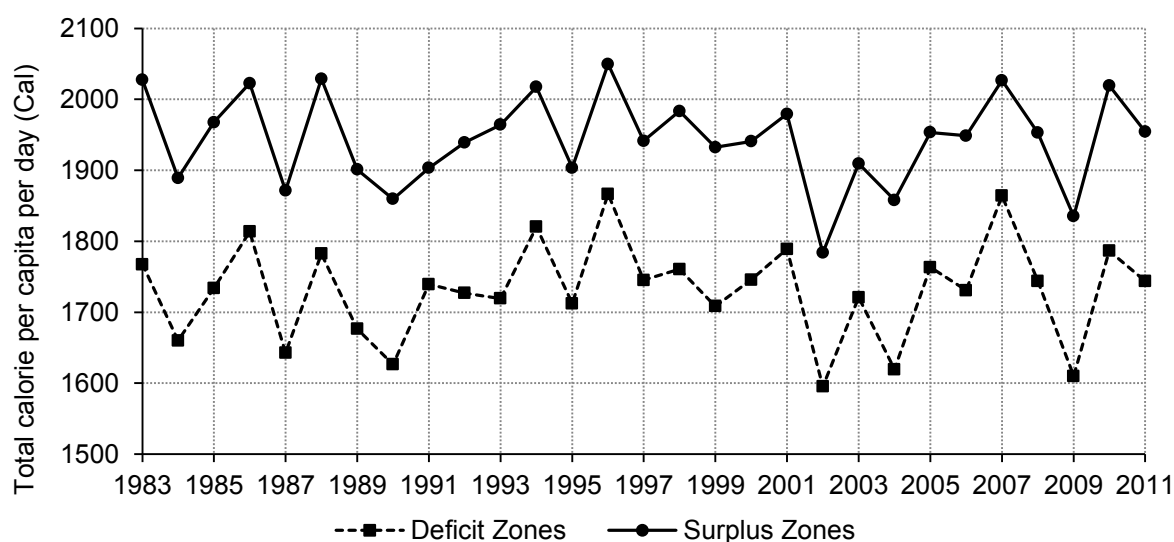


Figure 3-5: Total calorie per capita per day (Cal) averaged over food deficit zones and food surplus zones, respectively.

The poor population under a nationally defined poverty line (Diao et al., 2005) is even more strongly correlated with precipitation than GDP and calories per capita per day ($corr. = -0.73$, $p\text{-value} < 0.0001$; Table 3-1). Poverty rates in rural areas reach 55% and 54.3% in 1987 and 2002, respectively, approximately 10% higher than poverty rates in years with sufficient precipitation (e.g. 2007; Figure 3-6). As expected, urban poverty rates, ranging from 23.9% to 29.8%, are much lower than rural rates. The variability of poverty rates in urban zones is also lower than those in rural zones, as the poverty rate in urban zones is less affected by the climate conditions relative to the rural zones, even though they vary in the same direction (Figure 3-6). The total poverty rate is close to the rural poverty rate due to a large proportion of the population living in the rural area.

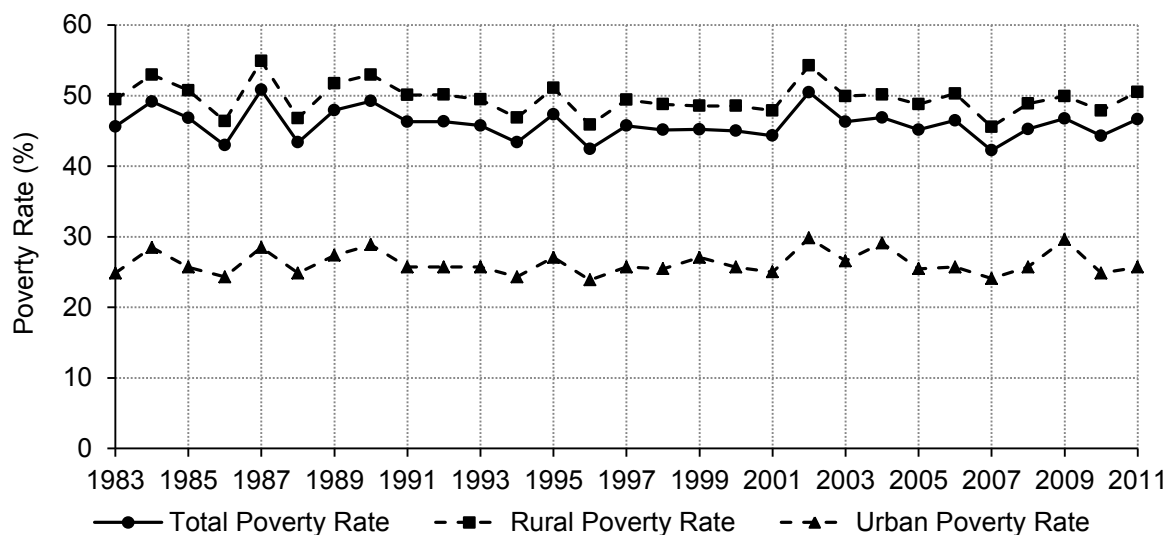


Figure 3-6: Percent of poor population over total population in each category (rural, urban, and total) over the simulated years.

3.3.3 Reallocating Teff and Maize Crop Land Simulations

Reallocating from teff to maize generates positive GDP benefits in most years with above-normal conditions, with mixed results for near- and below-normal conditions, but generally having reduced net benefit as precipitation drops (Figure 3-7a). In most of the below-normal years and some of near-normal years, the net benefit becomes negative, indicating in those years there is insufficient precipitation for maize to generate greater value than teff and therefore it is unwise to reallocate from teff to maize. Interestingly, when reallocating from maize to teff, no positive GDP benefit is apparent in any year (Figure 3-7b). While spatial variability could play a role – there might be positive net benefits for zones with extremely dry conditions and thus it becomes beneficial to grow teff instead, overall teff have reached a saturation point relative to maize, whose value has not been fully taken advantage of, particularly for years with above-normal conditions. In other words, aggregated across all zones, the current allocation of teff and maize may be suboptimal in all years.

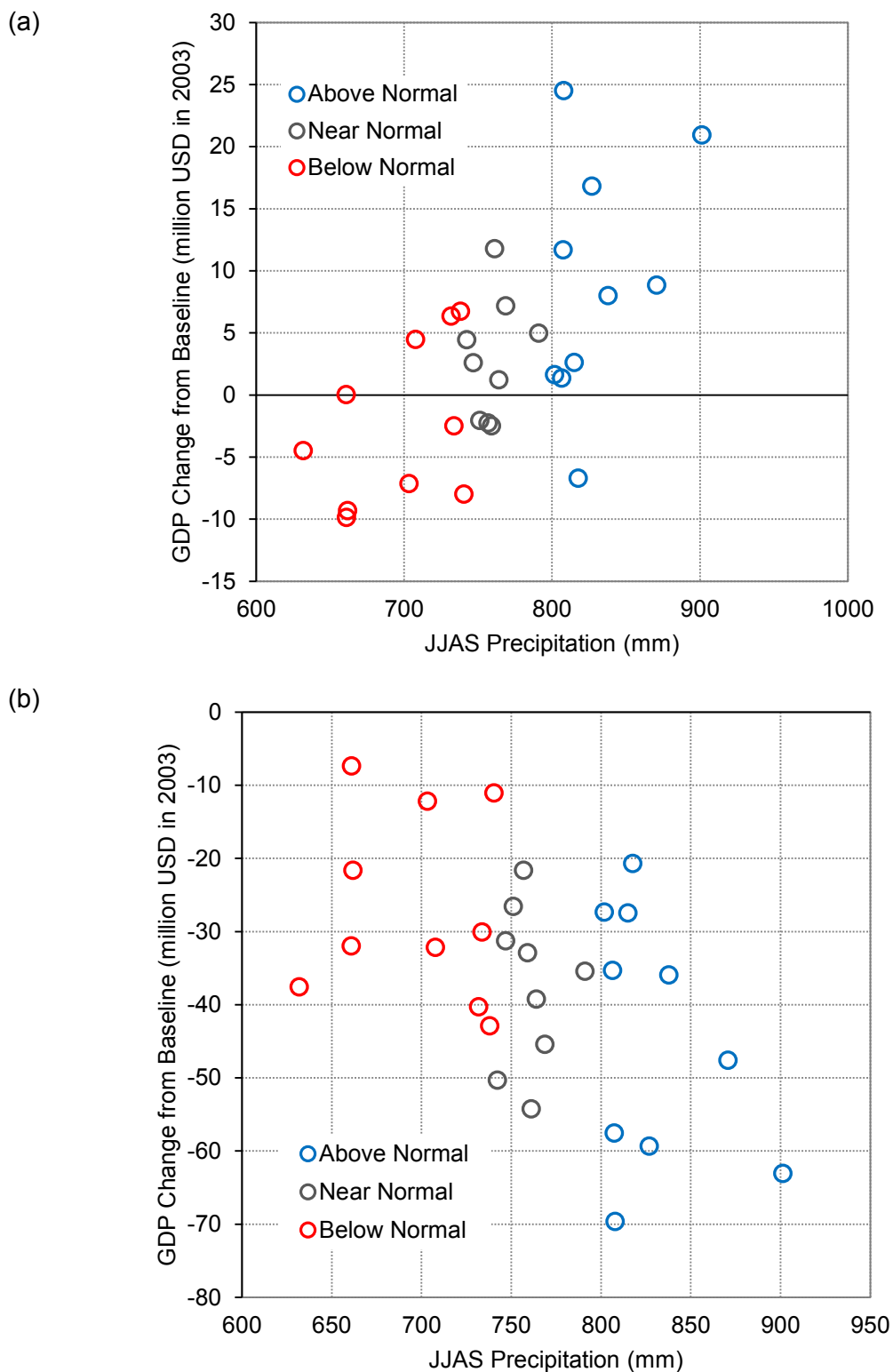


Figure 3-7: Real GDP change from baseline for all years with above-normal, near-normal, and below-normal conditions under (a) teff-to-maize and (b) maize-to-teff scenarios, averaged over outputs with different area reallocating percentages (5%, 10%, ..., 95%)

Hence, if the years with above-normal precipitation can be captured by prediction, it would be ideal to reallocate from teff to maize. In the seasonal prediction of JJAS precipitation over the 29 years (Figure 3-8), all the predicted above-normal years are correct (1988, 1996, 1998, 2003, 2007, and 2010), although some wet years are missed (1986, 1999, 2001, and 2006). Therefore, for this simulation, in the six years with predicted above-normal conditions, the planting area is reallocated from teff to maize over the full range of reallocating percentages.

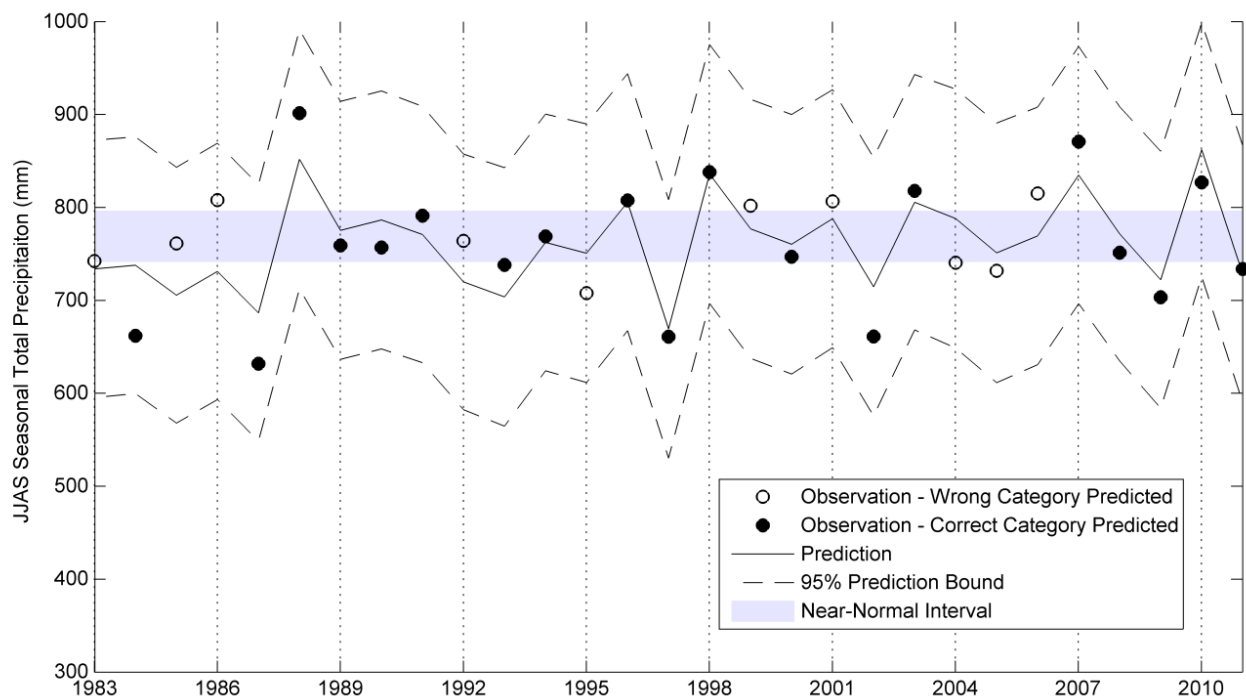
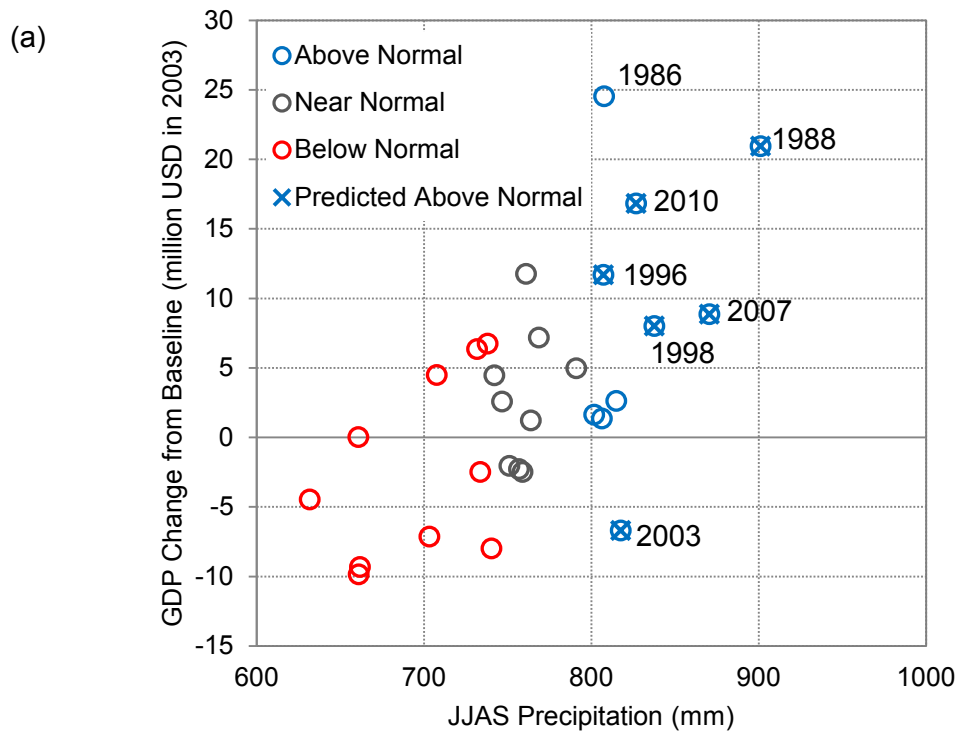


Figure 3-8: Seasonal prediction of JJAS precipitation from 1983 to 2011.

As a result, average positive change in GDP from the baseline occurs in five of the six years, except 2003 when the GDP decreases 6.7 million USD on average (Figure 3-9a). This is mainly due to heterogeneous distribution of precipitation across zones, and zones with large maize areas actually receiving less precipitation than that under normal conditions, even though the spatially aggregated precipitation is high in 2003. Even 5% reallocating from teff to maize in 2003 results

in a 0.22 million USD reduction in GDP (Figure 3-9b). In other predicted above-normal years, the 95% reallocating option results in the largest annual net benefit, ranging from 11.3 to 34.9 million USD per year (Figure 3-9b). For those wet years the prediction model missed, lost benefits are generally small, except 1986 when the average net benefits would have reached 24.5 million USD - the highest among all years.



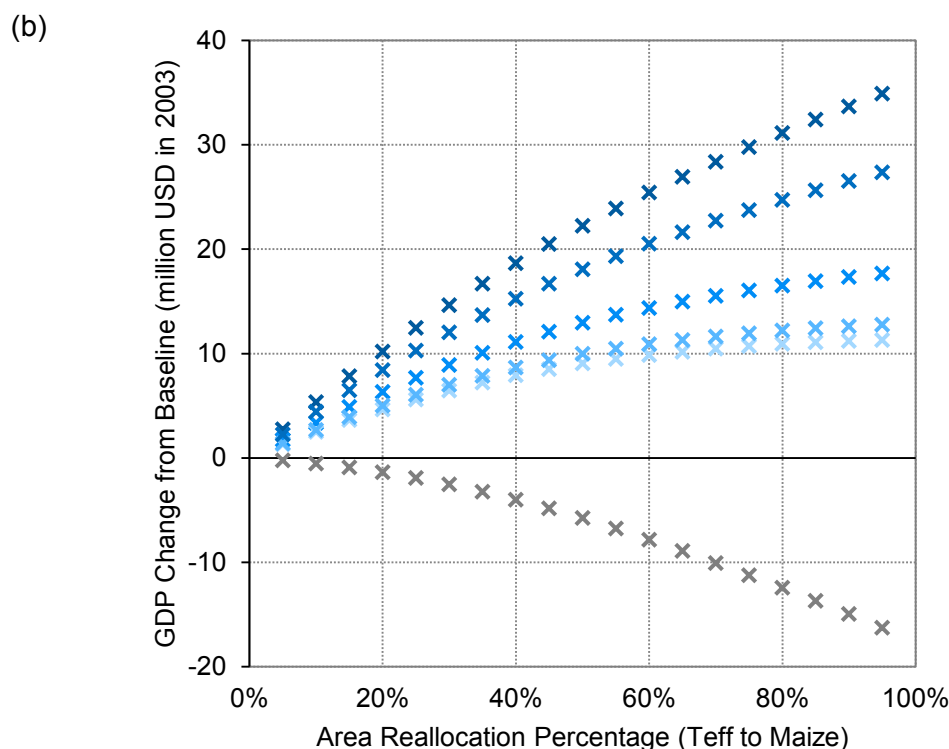


Figure 3-9: Real GDP change from baseline under “teff-to-maize” scenario (a) averaged over outputs with different area reallocating percentages (5%, 10%, ..., 95%), and (b) for each reallocating percentage respectively. *Note that years with above-normal condition predicted are marked with crosses (prediction-guided outputs).*

The average calories per capita per day for above-normal predicted years under the 95% area reallocation also increases from the baseline up to a maximum of additional 86 Cal based on all food consumption (Figure 3-10). As maize and teff are the major staple crops in Ethiopia, their total effect on caloric increase is close to all food consumption. Calories from teff consumption decreases gradually due to a lower supply as more maize is planted instead, however the calorie change stops at -40 Cal at a reallocating percentage of 55% (Figure 3-10). This is because teff starts to be imported from other countries at 55% teff-to-maize area reallocation due to a decrease in teff supply and a drastic increase in its domestic price (Figure 3-11). In 2003, the import quantity of teff is the lowest among the six above-normal years predicted, mainly due to a lower level of GDP in that year which in turn constrains the domestic food demand including teff.

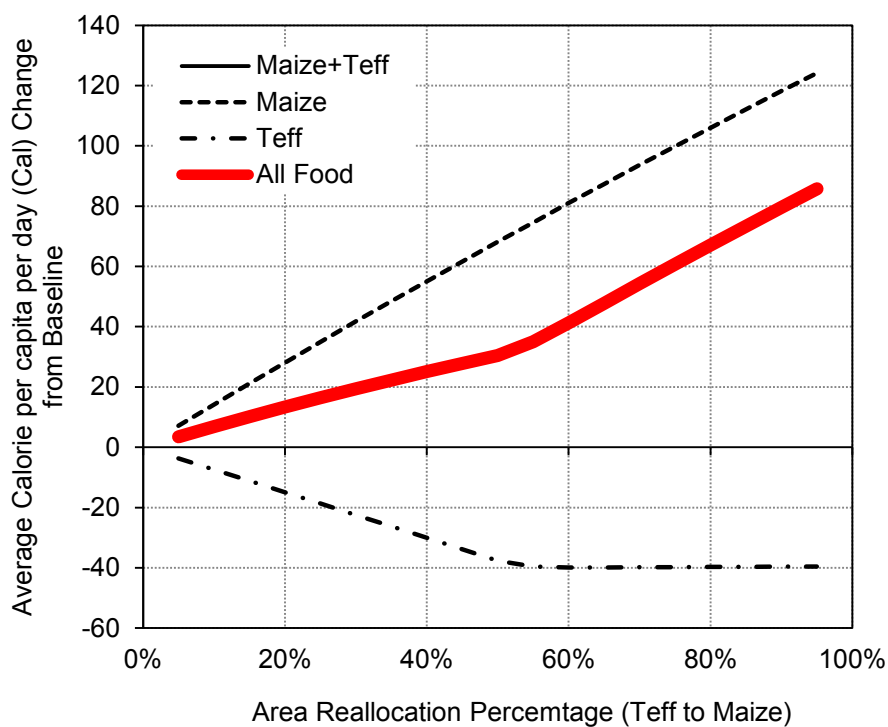


Figure 3-10: Calorie per capita per day (Cal) change from baseline averaged over predicted above-normal years for different teff-to-maize area reallocation percentages (prediction-guided outputs).

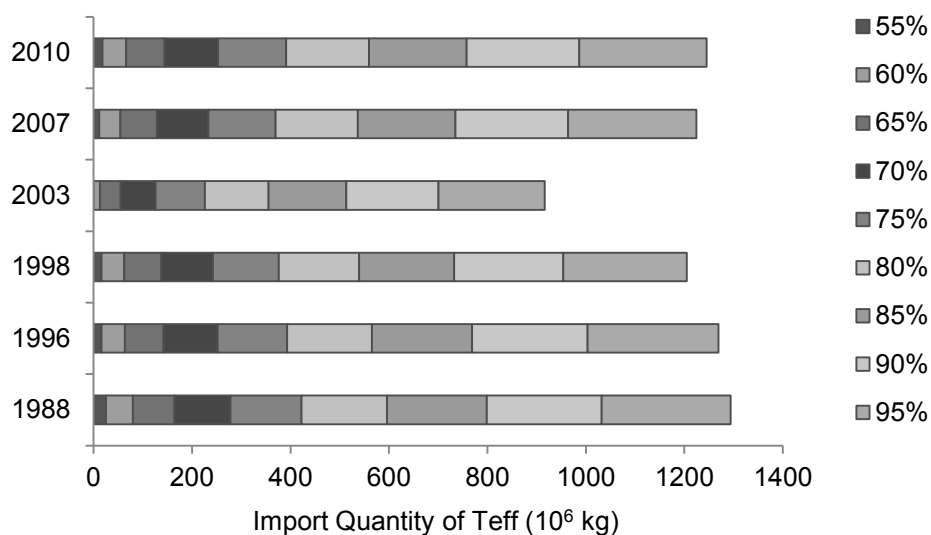


Figure 3-11: Import quantity of teff for predicted above-normal years across different teff-to-maize area reallocation percentages (import of teff is triggered starting at 55%).

The change in poor population and poverty rate illustrates a different picture than GDP and calories. Instead of best conditions occurring at an area reallocation percentage of 95%, they occur at 25% – 154,000 reduction in rural poor population, 12,000 reduction in urban poor population, and 0.25% reduction in overall poverty rate. For a 95% value, the rural poverty actually increases, although urban poverty decreases, which results in an approximately zero change in overall poverty rate (Figure 3-12). This highlights the uneven welfare distribution overall and should be considered when adopting policies.

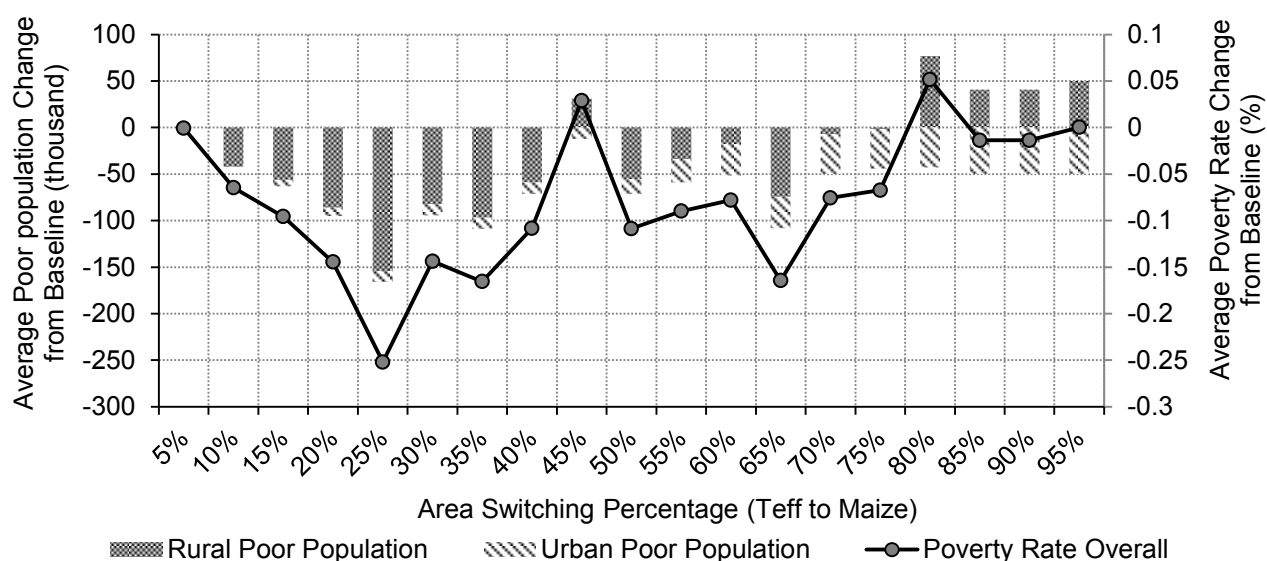


Figure 3-12: Poor population change (rural and urban) and poverty rate change from baseline averaged over action-adopted years.

3.4 Conclusions and Discussions

In this chapter, climate variability is imposed on an agro-economic model to simulate fluctuations in Ethiopia's economy. With a detailed agricultural structure embedded in the model, the influence on economic indices such as GDP, calorie consumed, poverty rate, etc. given climate variability, *ceteris paribus*, is well captured. Climate clearly plays an important role in

Ethiopia's economy with poverty rate being the most sensitive to total precipitation of the main rainy season. This is consistent with the findings in Diao and Pratt (2007) that growth in staple crops contributes strongly to poverty reduction. Seasonal prediction of precipitation is then applied to investigate potential net benefits compared to the baseline. One farmer action – reallocating between crop types on a fixed planting area – is tested here. As a result, in the six years with predicted above-normal precipitation, reallocating from teff to maize results in an average of 9.9 million USD (2003 currency) additional profits (positive in five of the six years, averaged over all percentage levels). Calorie consumption also increases, indicating that in above-normal precipitation years, both the maize value and its calories are higher than teff per unit of planting area. However, poverty rate does not always decrease with increasing GDP, revealing an uneven distribution of welfare. It is also worth noting that the higher reallocation percentages such as 95% may not be realistic, considering the access and availability of seeds and its drastic effect on price fluctuation. Additional analysis is required from a social-economic perspective to eventually make a practical policy in response to climatic prediction.

The net benefit attributable to prediction in this chapter is a relatively conservative estimate, given only one prediction-guided option is investigated. In the future, reallocating between other crop types could also be explored. Additional scenarios such as applying fertilizer or using improved seed according to the predicted climate conditions are also promising.

Spatial heterogeneity may also influence the prediction-guided benefit. In this chapter, given the initial efforts in exploring possible beneficial actions using prediction, understanding the economic effects brought by prescribed actions, and evaluating the prediction overall, no spatial variability in the adopted actions is considered. Not surprisingly, precipitation categories differ

across zones in any given year. If spatial variability is explicitly considered, actions for each zone could also be unique; varying actions across zones could potentially increase the overall benefit if predictions are skillful.

Allowing for dynamic reallocation across years may also bolster benefits, for example, using the prediction tendency to determine the extent of action taken (e.g. percent of area to reallocate). Moreover, adjusting prediction categorical intervals may be beneficial as some of the defined near-normal years also indicate positive net benefits when reallocating from teff to maize. Applying other formats of predictive information is also possible. However, the purpose of this chapter is not to optimize the value of using prediction, rather to present an innovative way to evaluate predictive information using economic indices at country level based on prediction-guided actions, which can provide a foundation for communications, decision-making, and strategic planning.

Chapter 4 Ethiopia's Grand Renaissance Dam: Implications for Downstream Riparian Countries

4.1 Introduction and Background

The Nile Basin, boasting the world's longest river, comprises numerous distinct hydroclimatic regions (Figure 4-1). The southernmost region includes a series of interconnected lakes in Tanzania, Kenya, and Uganda. This feeds the Sudd region in South Sudan, a vast series of swamps and lagoons that exhibit a regulating effect by expanding greatly during times of high inflow, allowing for increased evaporation, and releasing at a more moderate rate (Sutcliffe and Parks, 1987). Two distinct rainy seasons are evident throughout much of the basin due to the annual cycle of the Intertropical Convergence Zone (Block and Rajagopalan, 2009). Streamflow exiting the swamps converges at Malakal, in northeastern South Sudan, and continues northward through the fertile Sudanese central plains to Khartoum, Sudan, joining the Blue Nile River.

The Blue Nile River begins at Lake Tana in the northwestern Ethiopia highlands and is joined by many tributaries prior to reaching the Sudanese border. Precipitation in the highlands is concentrated in the northern hemispheric summer season, with nearly three-quarters typically falling between June-September (Conway, 2000); the annual average is approximately 1,300 mm with significant interannual variations. Once in Sudan, the Blue Nile continues flowing northwest, and is joined most notably with the Dinder-Rahad River, together flowing through the Gezira Scheme – one of the largest irrigation project in the world – and then converges with the White Nile. From Khartoum, the main Nile flows north, entering the Sahara Desert, and is joined from the east by the Atbara River approximately 300 km north of Khartoum. The Nile then

enters Lake Nasser on the Sudanese-Egyptian border, created by the High Aswan Dam, and subsequently flows through Egypt to the Mediterranean Sea.



Figure 4-1: The Nile basin with the forthcoming GERD. Base map courtesy of World Wildlife Fund (WWF, 2007)

Although the White Nile Basin size is more than five times that of the Blue Nile Basin, the latter contributes significantly more streamflow to the main Nile River. The total streamflow emanating from the Ethiopian highlands through the Blue Nile and Atbara Rivers accounts for approximately 84% of the inflow to Lake Nasser at Aswan, Egypt (Hurst, 1952; Figure 4-2; Block and Strzepek, 2010). The Blue Nile itself contributes approximately 65% of the total Nile flow into Lake Nasser (Yates and Strzepek, 1998).

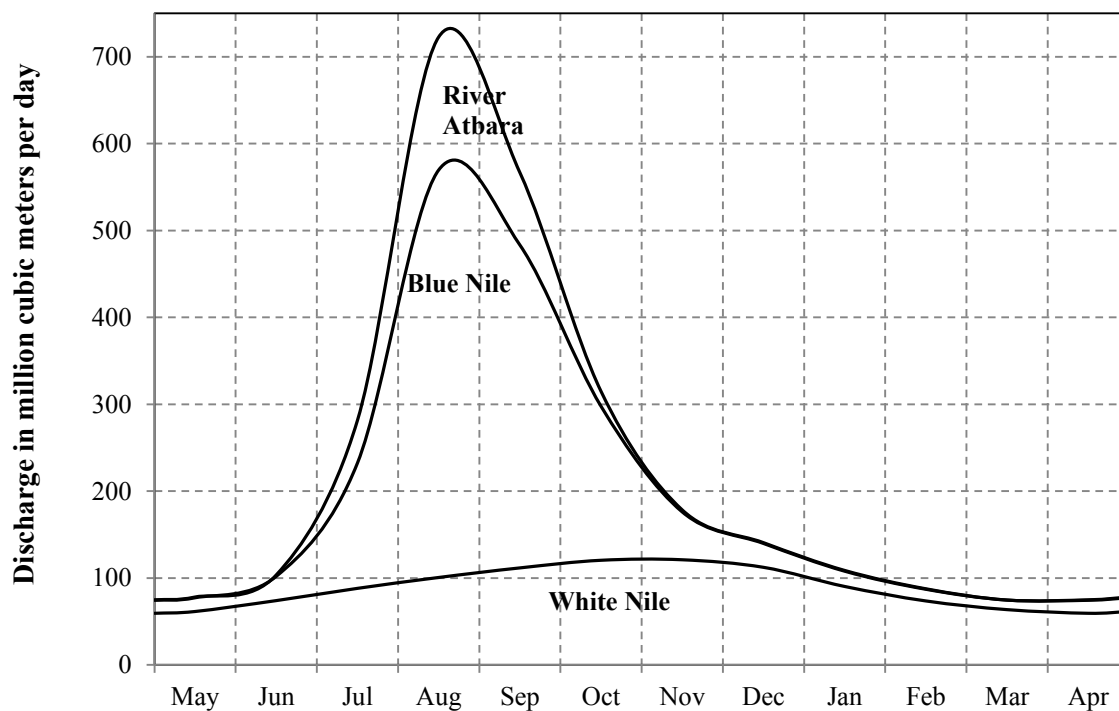


Figure 4-2: Contributing rivers to the main Nile discharge at Lake Nasser using monthly discharge data from 1961-1985 (for Atbara River and Blue Nile) and 1961-1995 (for White Nile) (Adapted from Hurst 1952).

The Grand Ethiopian Renaissance Dam (GERD) is an ambitious hydropower project that is being constructed on the Blue Nile, close to the border between Ethiopia and Sudan (Figure 4-1). It is reported that it will have a reservoir capacity of 74 km³ and, when fully operational, will have a power capacity of 6,000 Megawatts (MW). Construction began in 2011, and it is due to be commissioned in 2017 (EEPCo, 2013).

Ethiopia is endowed with significant water resources, providing an estimated hydropower potential of 45,000 MW nationally (Salman, 2013). Seeking sustainable economic growth, Ethiopia aims to exploit this potential, and is undertaking a larger program of dam construction on several of its rivers (Block and Strzepek, 2012).

There are several reasons behind Ethiopia's motivation to develop reservoir capacity and hydropower. First, the climate and hydrology of Ethiopia is extremely variable, on both inter- and intra-annual time scales, and floods and droughts are a recurring problem. A strong correlation between annual rainfall and the gross domestic product (GDP) growth rate has been demonstrated (WorldBank, 2006), and as the majority of Ethiopians rely on rain-fed agriculture, this variability is closely linked to vulnerability. Therefore the development of storage may partially buffer the effects of such climate extremes.

Second, Ethiopia's hydropower capacity has been described as a 'blue gold' that could contribute to economic growth, and a reliable source of electricity (Verhoeven, 2011). Figures from the World Bank show that only 17% of Ethiopians have access to electricity (WorldBank, 2013). Its population has grown rapidly, rising from 66 to 87 million between 2000 and 2010, and it is projected to reach 110 million by 2020 (UnitedNations, 2013). Demand for electricity has been growing accordingly, and is expected to continue to grow. Furthermore the export of electricity to its neighbors could promote regional economic growth and cooperation, as well as strengthening Ethiopia's position as a regional power (Verhoeven, 2011).

Ethiopia's downstream neighbors, most notably Egypt, however, have voiced their concern over the construction of the GERD. The Nile River is an international watercourse shared by 11

countries and has historically been a source of tension (Waterbury, 2002). A series of treaties have given Egypt and to a lesser extent Sudan a position that has been described as “hydro-hegemony” (Zeitoun and Allan, 2008). A treaty made in 1929 between Egypt and the United Kingdom allocated 48 km³ of the utilizable flow to Egypt and 4 km³ to the Sudan. Egypt also assumed the right to veto any construction projects in the upper riparian nations that would harm its interests. In 1959, in preparation for the construction of the High Aswan Dam, Egypt and an independent Sudan signed the Agreement for the Full Utilization of the Nile Waters, which allocated 55.5 and 18.5 km³ to Egypt and Sudan respectively, with 10 km³ of a total utilizable flow of 84 km³ lost as evaporation and seepage (Conway, 2005). This agreement thus allocated all of the utilizable flow, as measured at the Aswan Dam, to Egypt and Sudan exclusively.

The upstream Nile Basin nations, including Ethiopia, contest these treaties (Waterbury, 2002). In 1999, the Nile Basin Initiative was launched to promote cooperation between the nation states. It is an intergovernmental organization partnership run by all the riparian states (except Eritrea who is an observer), at a ministerial level, supported by International donors, including the World Bank (Reuters, 2013). A direct outcome of the Nile Basin Initiative, the Cooperative Framework Agreement (CFA), was signed by Ethiopia, Kenya, Uganda, Rwanda, and Tanzania in 2010, and Burundi in 2011 (Salman, 2013). The CFA establishes each country’s right to an “equitable and reasonable utilization” of the Nile Waters, as well as the obligation to avoid significant harm to its riparian neighbors. In June 2013, the agreement was ratified in the Ethiopian parliament, and the newly independent South Sudan and the Democratic Republic of Congo are expected to sign the agreement in the near future (WashingtonPost, 2013). Egypt and Sudan have opted not to sign the CFA, highlighting the tensions that existed even before the announcement of the GERD.

Egypt's primary concern regarding the construction of the GERD is ensuring access to water resources from the Nile – its primary source – to serve its livelihoods and rapidly growing population; national water demand is anticipated to increase in the future (Swain, 2011). In order to investigate the potential impacts of the dam, a 10 person international commission was convened, and submitted its findings in May 2013. Although not yet publicly available, the President of Egypt has been reported as referring to the studies conducted so far as “not adequate” (Reuters, 2013). Sudan, on the other hand, has expressed their support for the GERD project, as dam operation may help to increase power supply at a competitive rate and significantly reduce the silt and sediment downstream. Additionally, the cost to remove mud from the irrigation channel of the Gezira Scheme would be largely reduced (MoFA, 2012). However, the streamflow quantity available for the Gezira Scheme may be reduced, particularly during the filling stage.

The GERD will inevitably alter the hydrology of the Nile Basin, particularly during the reservoir filling stage, but also under normal operation. The reservoir filling rate policy, climate variability, and climate change all represent important drivers in the GERD's performance and – of principal interest in this chapter – on downstream flows. Characterizing the uncertainty attached to each of these drivers is also imperative for properly understanding the potential range of effects. The motivation behind this chapter is to provide regional decision-makers with a set of plausible, justifiable, and comparable outcomes regarding how streamflow entering the Gezira Scheme and Lake Nasser post-construction of the GERD may vary, contingent on the aforementioned drivers.

4.2 Nile Basin Modeling Framework

4.2.1 Nile Hydrology and GERD Reservoir Models

The hydrology model adopted here is a variation of the Water Balance Model (WatBal), developed as an integrated tool for modeling the response of river basins to climate impacts (Yates, 1996). It requires inputs of monthly precipitation, daily mean temperature and temperature range, producing monthly streamflow within a river basin. WatBal is based on three main modeling components: the soil moisture modeling scheme, which accounts for the water flux in a conceptualized basin, potential evapotranspiration using Hargreaves method (Hargreaves and Samani, 1982), and a reservoir storage scheme for lakes and swaps. For the Nile Basin case, WatBal has been previously calibrated and shown to adequately reproduce observed monthly streamflow along the Nile River (Yates and Strzepek, 1998). The model links four sub-basins to represent the full basin (Figure 4-3). Losses between confluences are considered linear, based on historical streamflow (Yates and Strzepek, 1998).

Given the limited number of streamflow gauging stations along the Nile River, three important locations with available and sufficient streamflow station data are selected to recalibrate the hydrology model described above. These locations include 1) the Roseires dam located in Sudan, just downstream of the GERD on the Blue Nile, 2) Khartoum at the confluence of the White Nile flowing north and the Blue Nile flowing west from Ethiopia, and 3) Dongola just upstream of Lake Nasser.

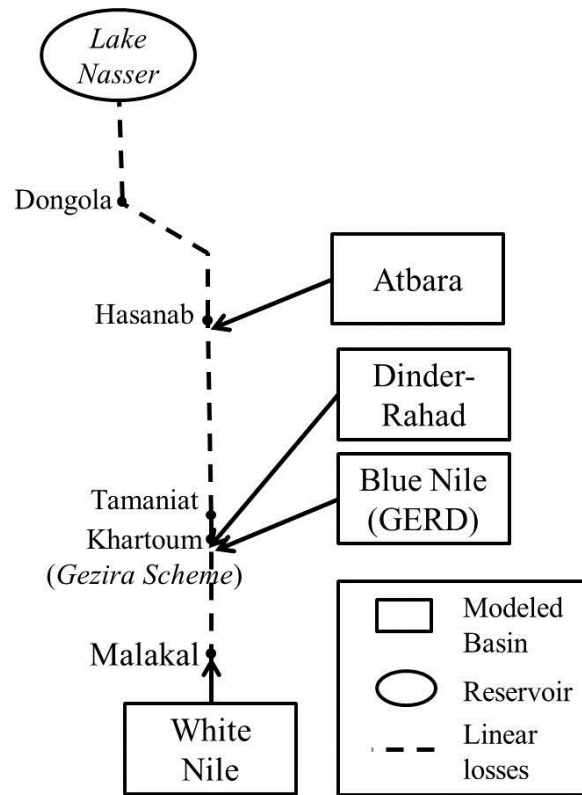
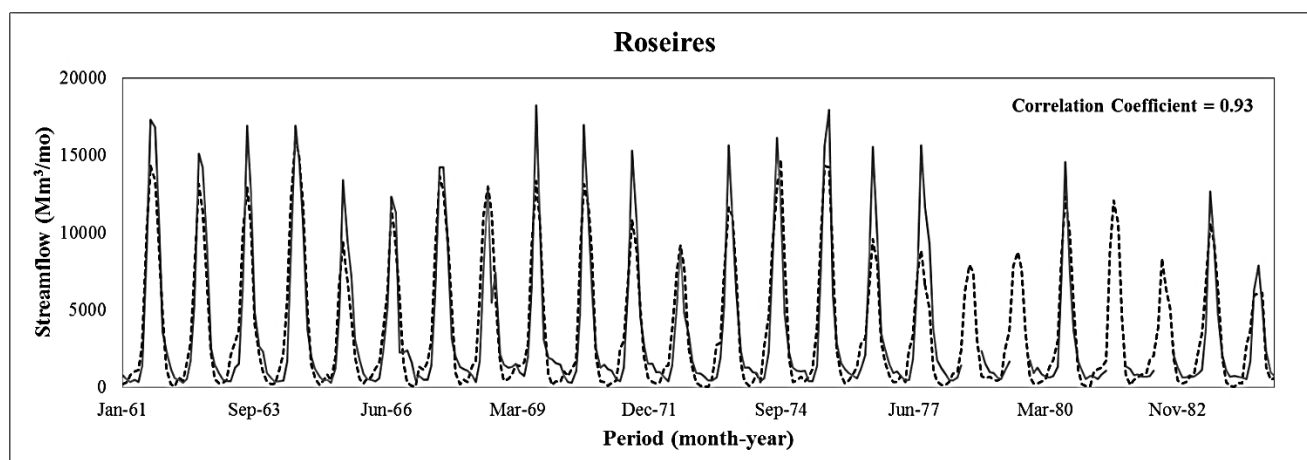


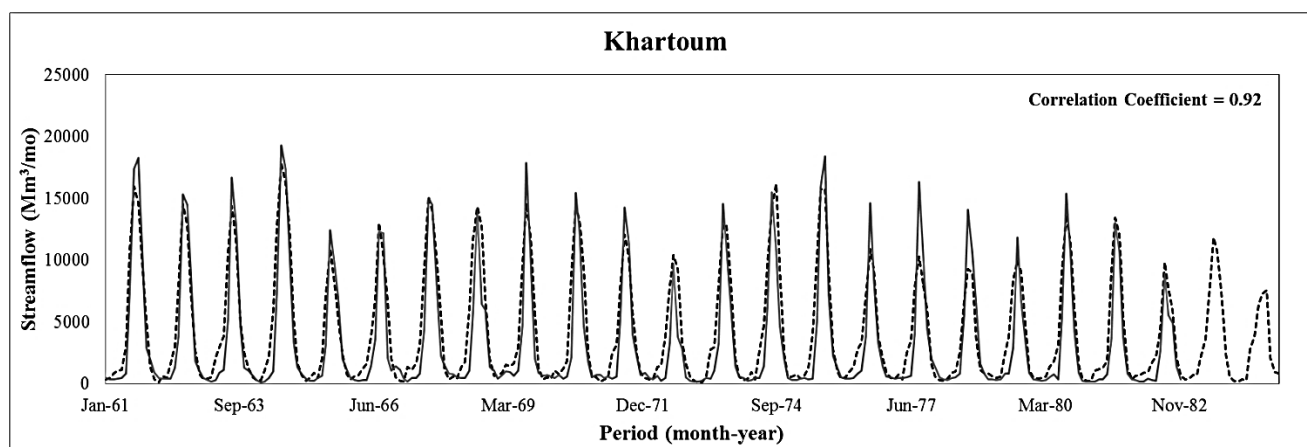
Figure 4-3: Schematic of the model framework

Calibration results at each location with estimated correlation coefficient between observed and modeled streamflow data are presented below (Figure 4-4a, b, and c). Each demonstrates a correlation coefficient greater than 0.9, indicating strong agreement. To assess the model error, cumulative errors in volume across months at Dongola are evaluated over 1961 to 1995 (Figure 4-4d). It indicates that at Dongola during low streamflow months from January to April our model underestimates the streamflow, followed by an overestimate until August. The overall cumulative result shows that we underestimate the average annual streamflow by 3,299 million m^3 . This is about 4.5% of the observed average annual streamflow. To minimize this influence and emphasize the differences between flow policies, we use percent difference between no dam scenarios and filling policies, with the assumption that model errors carry over accordingly and therefore nearly cancel. This is discussed further in the modeling results section.

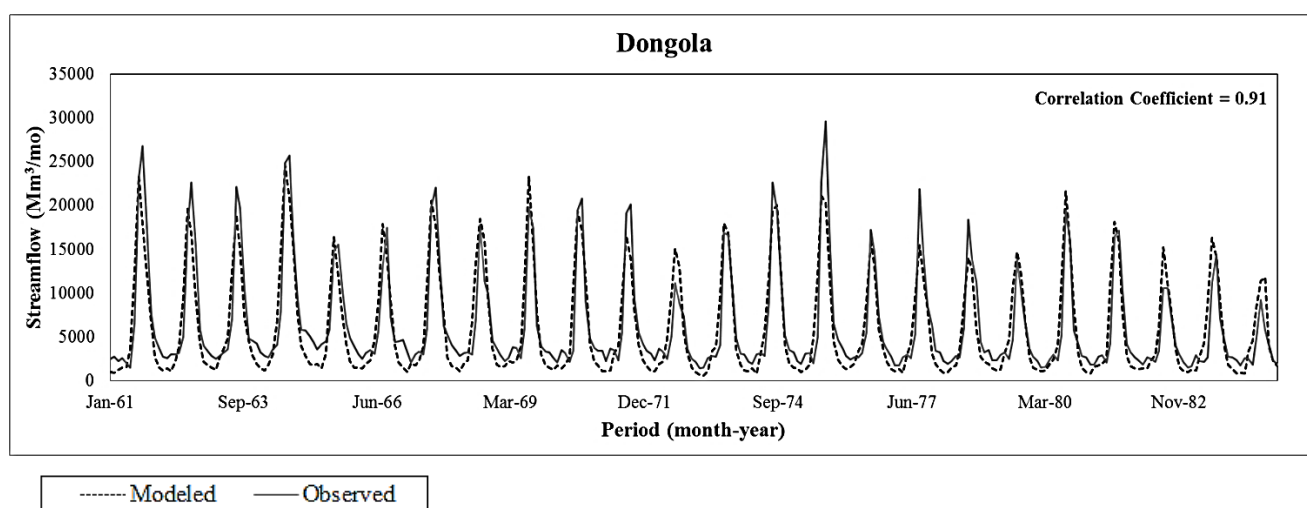
(a)



(b)



(c)



(d)

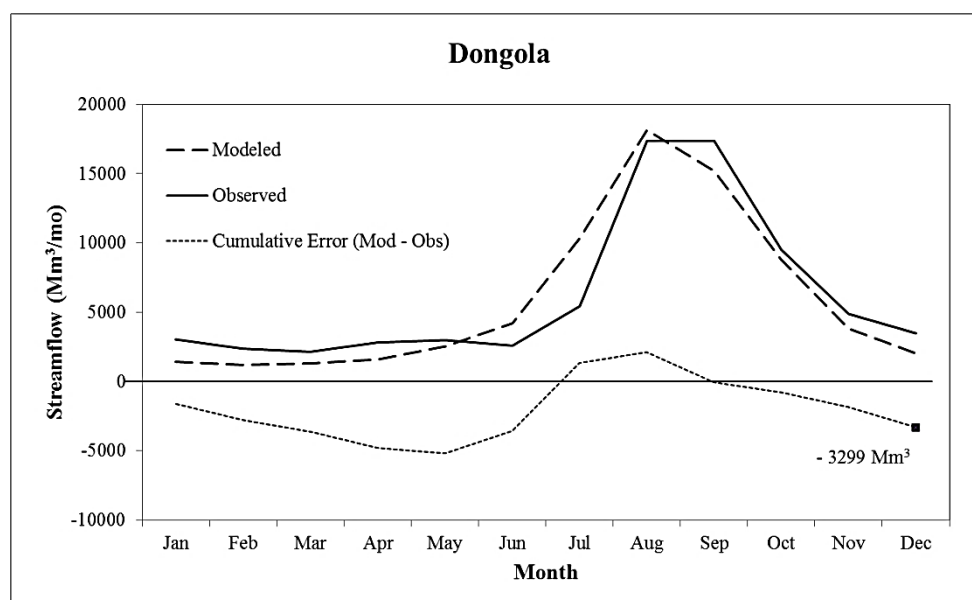


Figure 4-4: Nile WatBal hydrology model recalibration on monthly streamflow at (a) Roseires (b) Khartoum and (c) Dongola; (d) Monthly streamflow averaged over 1961-1995 and monthly cumulative error in volume of modeled and observed streamflow.

The GERD hydropower and reservoir model simulates monthly reservoir volume, hydropower generation, evaporation, and downstream flows, as well as the time required to reach the reservoir full supply level (FSL), based on outputs from the Blue Nile Basin hydrology component (King and Block, 2014). In this chapter, the downstream implications are further explored, given the non-linearities in streamflow reduction between the GERD and other points of interest in the basin – such as the Gezira irrigation scheme and Lake Nasser – to understand the effect of filling policies and climate variability from a probabilistic perspective. Plausible climate changes are also considered. The model is run over a 50-year period, representing 2011-2060. (Time series for the hydrology and reservoir models are discussed in section 4.2.4). Figure 4-5 illustrates the general methodological process. Hydropower production may commence once the minimal reservoir operating level is reached, however maximum generation is not possible until the FSL is attained. The FSL design volume is approximately 1.5 times the total annual

average streamflow entering the reservoir (Than, 2011). Reservoir surface area-to-volume and volume-to-head relationships were developed based on publicly available GERD design specifications and preliminary work done by the United States Bureau of Reclamation (USBR, 1964).

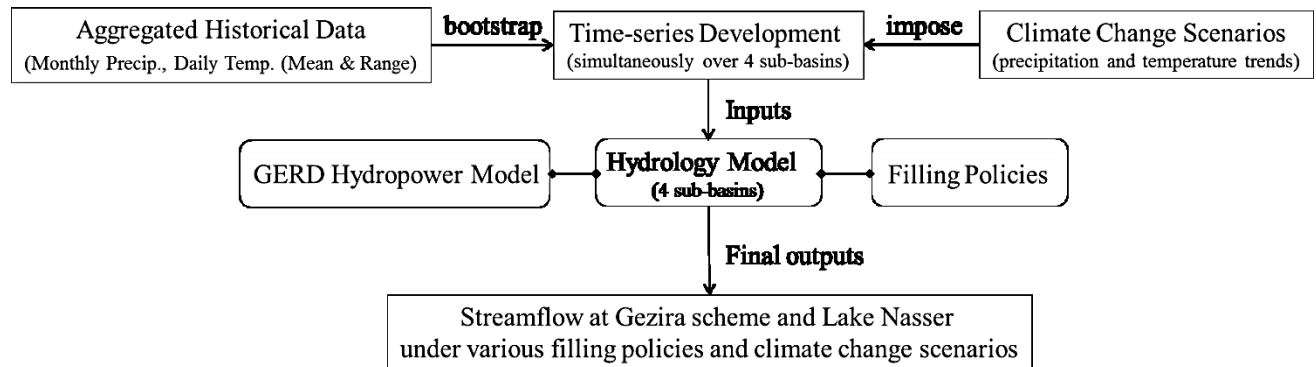


Figure 4-5: Flowchart of methodology

According to the Ethiopian Electric Power Corporation (EEPCo, 2013), the first phase of reservoir impoundment will commence after 44 months of construction (in 2014), allowing up to 9 million m³ (SudanVision, 2013). During this phase, two 375 MW turbines will be available for hydropower generation. Once full dam construction is completed (2017) impoundment may continue up to the FSL volume of 74 million m³, and all turbines will be available for hydropower generation, providing 6,000 MW of capacity.

The volume of water impounded in the reservoir is a function of existing storage, incoming streamflow, net evaporation, and reservoir releases (through the turbines or spilling). The rate of impoundment, or filling policy (discussed in Section 4.2.3), thus becomes critically important, balancing downstream flow reductions and hydropower generation, particularly in the early years prior to reaching the FSL.

4.2.2 Model Data Requirements

The hydrology and hydropower models require inputs of monthly precipitation, daily mean temperature, and the diurnal temperature range. The University of East Anglia Climatic Research Unit's (CRU) TS 3.20 dataset, on a 0.5×0.5 degree grid for 1951-2000, is utilized to provide all variables (Harris et al., 2014). This 50 year historical period is selectively chosen to minimize gridded data represented by sparse station data (prior to 1950) as well as minimize any possible effects of climate change (post 2000). Note that we shifted from NMA data to CRU data given that the NMA data was not available at the time when this work was performed. Streamflow data was acquired through the National Center for Atmospheric Research's ds552.1 dataset (Global Runoff Data Center et al., 2001).

4.2.3 GERD Filling Policies

Presently, there is no publicly available information regarding Ethiopia's reservoir filling plan. Historically, there are analogous situations, particularly related to filling policies and downstream impacts that would ideally be avoided. For instance, when the Turks started to fill the reservoir behind the Ataturk Dam on the Euphrates River in 1990, water to the downstream riparian countries – Syria and then Iraq – was restricted for a period of one month (Kibaroglu and Ünver, 2000). These actions, including insufficient discussions and agreements prior to reservoir impoundment, ignited a heightened conflict between all involved riparian countries. The United Nations Convention on Non-navigational Uses of International Watercourses, the most influential and frequently cited rules, provides general principles for sharing transboundary river water (Schwabach, 1998). Two of its major principles are “equitable utilization” and “no

significant harm”, that is, to utilize an international watercourse in an equitable and reasonable manner without significant harm to downstream riparian countries (Clemons, 2008). Clearly, the selected filling rate will have implications on the time to fill the reservoir, hydropower generation, and downstream flow reductions. Filling the reservoir more quickly likely favors Ethiopia, as hydropower generation could begin in earnest sooner; filling the reservoir more slowly would limit reductions to downstream flows and may lessen the impacts on downstream activities and livelihoods. To evaluate reasonable utilization of water for the GERD in Ethiopia and correspondingly its impacts on downstream riparian countries – Sudan and Egypt – five filling policies are explored: three considering a fractional retention of total monthly streamflow entering the reservoir, impounding 5%, 10%, or 25% of Blue Nile flow through the GERD; and two additional threshold-based policies with retention rates contingent on the historical average streamflow (HASF) (i.e. to retain any quantity greater than HASF, or any quantity greater than 90% of HASF). The fraction-based filling policies therefore guarantee available water for storage, however the threshold-based filling policies do not guarantee available water for storage in years of below-average streamflow. During these months of no retention under the threshold policies, the full volume released downstream can still be routed through the turbines for hydropower generation.

The volume of water (Vol) present in the reservoir during any month (t) is a function of the previous month’s volume, $Vol(t-1)$, the filling policy, the streamflow into the reservoir (Q_{in}), and the net evaporation (nET). These functions take a slightly different form for the fraction-based (Equation 4-1) and threshold-based (Equation 4-2) policies (King and Block, 2014):

$$Vol(t) = Vol(t - 1) + Fill\ Policy * Q_{in} - nET(t) \quad \dots\dots\dots (Equation\ 4-1)$$

$$Vol(t) = Vol(t - 1) + \max(Q_{in} - HASF(t), 0) - nET(t) \quad \dots\dots\dots \text{(Equation 4-2)}$$

4.2.4 Time series Development

To assess the performance of the GERD and implications on downstream flows, the hydrology and hydropower models are run over 2011-2060 – from beginning of construction to the estimated economic life of the dam. Of particular interest is the early decade(s) during which the reservoir is being filled and downstream flow reductions will be most notable. The 1951-2000 CRU data represents one plausible time series for hydrology model input, however to capture the uncertainty in possible year-to-year (climatic) variability, 100 50-year simulations are created by bootstrapping from the CRU precipitation, temperature, and daily temperature range data simultaneously to preserve the inter-variable relationships. Given strong month-to-month persistence in precipitation, bootstrapping was executed at the annual scale (full calendar years). Minimal autocorrelation is evident between years (December to January) in the historical record, and was therefore excluded. For consistency, bootstrapping was performed simultaneously across each hydrology model sub-basin (White Nile, Blue Nile, Atbara and Dinder-Rahad) to preserve spatial correlation.

Recent studies based on general circulation model (GCM) output illustrate expected long-term climate changes over Ethiopia and East Africa (Solomon et al., 2007; Giannini et al., 2008; Conway and Schipper, 2011), warranting inclusion into this analysis. Models consistently indicate an increase in future temperature, with an increase of 2.5 degrees Celsius (C) by 2060 as a middle-of-the-road estimate. The models are in less agreement, however, regarding the direction and magnitude of precipitation changes, ranging from modest decreases to substantial increases (Figure 4-6).

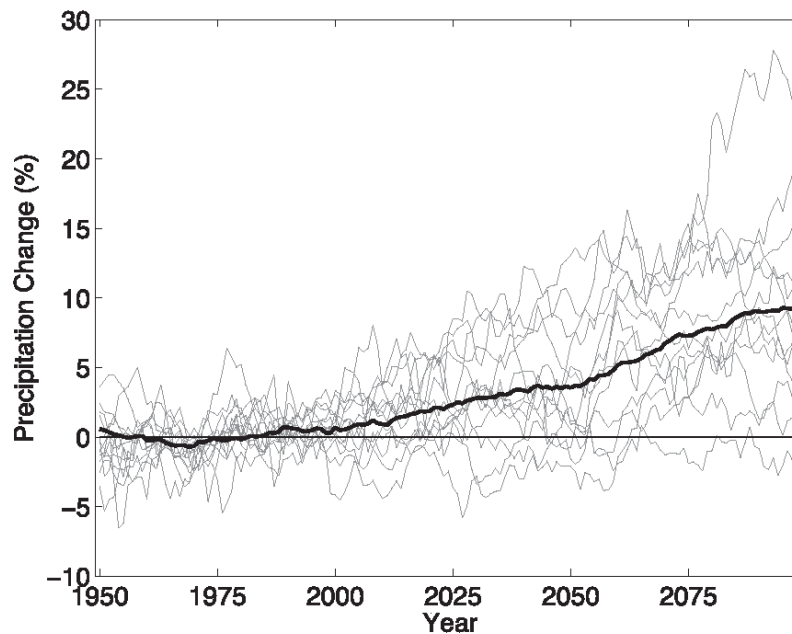


Figure 4-6: GCM simulations of historical and projected precipitation changes as a percent of the 1950-1990 average annual precipitation. The multi-model mean projection is represented by the bold line. GCMs are from the International Panel on Climate Change Fourth Assessment Report (IPCC 4AR) (Pachauri and Reisinger, 2007) based on the A1B scenario from 2000-2100 (Adapted from Giannini et al., 2008).

To impose potential climate changes, linear trends in temperature and precipitation changes are added to the 100 50-year time series previously constructed. Given consistent GCM projections of increasing in temperature in the future, a linear trend from no change (0-degrees C) in 2011 to +2.5 degrees C in 2060 is imposed. However, the suite of GCMs illustrates a large range of projected precipitation changes, thus a sensitivity-type approach is taken to better reflect what the GCMs are indicating as a group. The focus here is not to interpret individually how the GCMs may prescribe precipitation changes to evolve in the future but to perform a sensitivity analysis of plausible future changes based on the suite of GCMs. Accordingly, trends varying linearly from no precipitation change in 2011 to - 20%, -10%, -5%, 0%, +5%, +10%, +15%, and + 20% by year 2060 are added. No changes to future daily temperature range were imposed.

Thus, the 100 time series will be evaluated with 8 possible precipitation changes and 6 filling policies (including a no-filling policy or “no dam” option) to infer reductions in streamflow entering Gezira Scheme and Lake Nasser given reservoir filling and operation of the GERD.

4.3 Modeling Results for Streamflow at the Gezira Scheme and Lake Nasser

Downstream implications include a wide range of benefits, such as silt reduction (longer life for downstream dams), regulated flow for irrigated area, downstream hydropower generation, whereas the major loss is the streamflow reduction, which could lead to other potential impacts. In this chapter, reduction in annual streamflow downstream of the GERD is focused and discussed, as a product of the reservoir filling policy, net evaporation, climate variability, and climate change. The filling policy, net evaporation, and climate variability dominate reductions in the early years and decades prior to reaching the FSL; climate change and net evaporation (approximately 2 km³ per year) dominate once the FSL has been attained and into the future. Clearly, a reservoir filling policy impounding water at a greater rate will cause larger downstream flow reductions, but for a shorter duration. Under future conditions assuming no changes in precipitation, streamflow reductions into Lake Nasser attributable to reservoir filling may vary from seven years (on average) under the 25% flow policy to > 50 years under the 5% flow policy (Figure 4-7; the 5% filling policy results are not presented given their unlikelihood of adoption). The boxplot of reduction in annual streamflow at Lake Nasser in Figure 4-7 indicates consistently low reductions under the 10% filling policy, large reductions during the first few years and then decreasing beginning around year 2023 under the 25% filling policy, and large variance associated with >HASF filling policy. In the years that are significantly wetter than average, substantial reservoir impoundment is allowed, as can be inferred from the large

reductions under the HASF policy (outliers). Although the percent reductions in streamflow appear large in those wetter years for the HASF policy, the actual streamflow leaving the GERD is close to the long-term historical average. That is to say, the HASF filling policy allows up to the historical average flow to pass downstream, but produces an inconsistent impounding rate behind the GERD. The 25% filling policy allows the GERD to be filled relatively quickly, a security for Ethiopia, but for downstream users the flows are fluctuated and can be scarce during the dry years. Similar patterns can be found for the Gezira Scheme, but with larger average reductions in annual streamflow. This is a result of less tributary contributions downstream of the GERD before reaching the Gezira Scheme, as the White Nile River and the Atbara River have not yet joined.

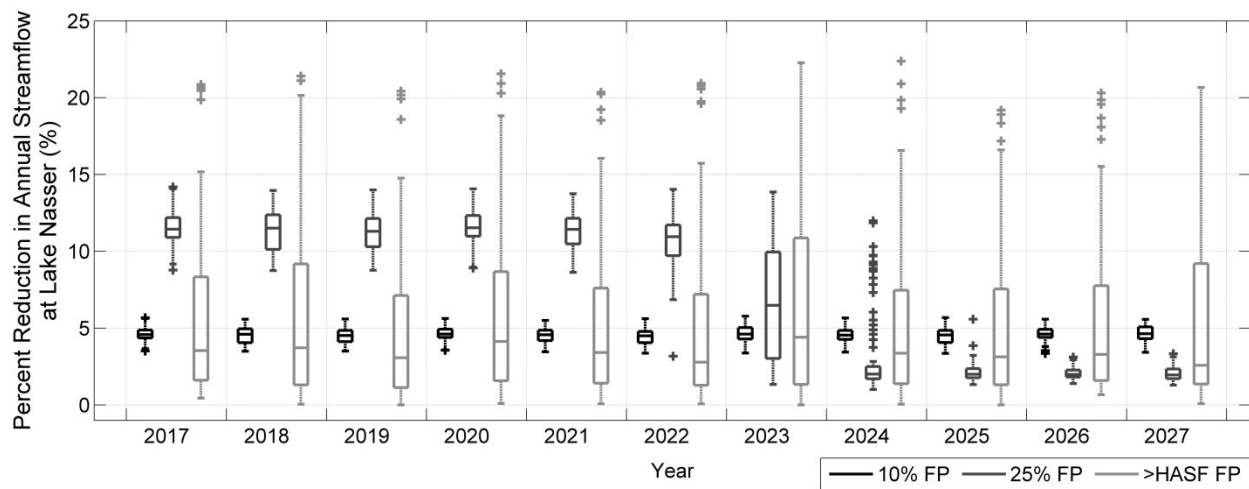


Figure 4-7: Percent reduction in annual streamflow entering Lake Nasser considering no future changes in precipitation. Three filling policies shown. *Note: For each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually in crosses.*

Defining the filling stage as 15 years (2017-2032), an average filling period across different filling policies consistent with King and Block (2014), it is possible to compare annual average streamflow during the filling stage at the Gezira Scheme and Lake Nasser across filling policies

and changes in future precipitation, relative to no dam, no precipitation change situation. All simulations do, however, account for a uniform increase in temperature. Considering median values (Table 4-1 and 4-2) and no precipitation trend, the percent change in annual average streamflow at the Gezira Scheme across filling policies ranges from -9.0% to -11.5%, while at Lake Nasser anticipated reductions are smaller, ranging from -4.6% to -5.9%. The 25% filling policy produces the largest reduction in streamflow at both locations, however not substantially different than for the threshold-based policies, especially under a wetter future climate. If the future becomes significantly wetter (+20% changes in precipitation), the additional streamflow from the four sub-basins will more than account for the volume impounded annually in the GERD reservoir. Under a drier future, streamflow reductions and the period required to fill the reservoir both increase.

Table 4-1: Median of 2017-2032 percent change in annual average streamflow at Gezira Irrigation Scheme for each filling policy and precipitation trend. All values are relative to the no dam policy and no changes to future precipitation.

<i>filling policy</i>	<i>Precipitation Trend, temp all +2.5 C</i>							
	-20%	-10%	-5%	0	+5%	+10%	+15%	+20%
10%	-18.0%	-13.6%	-11.2%	-9.0%	-6.5%	-4.1%	-1.6%	1.0%
25%	-21.7%	-16.7%	-14.0%	-11.5%	-8.7%	-6.0%	-3.3%	-0.3%
>HASF	-16.8%	-13.7%	-11.9%	-10.1%	-7.8%	-5.4%	-2.8%	0.0%
>0.9*HASF	-19.7%	-15.8%	-13.5%	-11.1%	-8.5%	-5.8%	-3.1%	-0.2%
No Dam	-10.0%	-5.1%	-2.4%	0.0%	2.7%	5.4%	8.1%	11.0%

Table 4-2: Median of 2017-2032 percent change in annual average streamflow at Lake Nasser for each filling policy and precipitation trend. All values are relative to the no dam policy and no changes to future precipitation.

<i>filling policy</i>	<i>Precipitation Trend, temp all +2.5 C</i>							
	-20%	-10%	-5%	0	+5%	+10%	+15%	+20%
10%	-11.2%	-7.9%	-6.2%	-4.6%	-2.9%	-1.2%	0.6%	2.4%
25%	-13.1%	-9.5%	-7.6%	-5.9%	-4.0%	-2.1%	-0.3%	1.7%
>HASF	-10.6%	-8.0%	-6.6%	-5.2%	-3.5%	-1.8%	-0.1%	1.9%

>0.9*HASF	-12.1%	-9.1%	-7.4%	-5.7%	-3.9%	-2.0%	-0.2%	1.8%
No Dam	-7.1%	-3.6%	-1.7%	0.0%	1.9%	3.7%	5.6%	7.5%

While useful for comparing filling policies and future climate changes, the defined filling period may be misleading. For example, annual average streamflow at the Gezira Scheme (Lake Nasser) under the 25% filling policy is only 1.3%-3.7% (0.7-1.9%) less than under the 10% policy. This is predominantly attributable to the time required to fill the reservoir; under the 25% policy, the reservoir fills relatively quickly (e.g. Figure 4-7) and almost always prior to 2032, thus the annual average across 2017-2032 includes years after the FSL has been attained, decreasing the overall percent reduction in annual average streamflow. Therefore, a shorter period (2017-2022) for comparison may provide more realistic differences (Table 4-3 and 4-4), although averaging over so few years can be highly influenced by an anonymously wet or dry year.

Table 4-3: Median of 2017-2022 percent change in annual average streamflow at Gezira Irrigation Scheme for each filling policy and precipitation trend. All values are relative to the no dam policy and no changes to future precipitation.

<i>filling policy</i>	<i>Precipitation Trend, temp all +2.5 C</i>							
	-20%	-10%	-5%	0	+5%	+10%	+15%	+20%
10%	-15.4%	-12.7%	-11.0%	-9.6%	-8.1%	-6.5%	-4.9%	-3.3%
25%	-27.8%	-25.5%	-24.0%	-22.8%	-21.4%	-19.9%	-18.3%	-16.8%
>HASF	-15.0%	-13.6%	-12.8%	-12.0%	-11.3%	-10.6%	-10.0%	-9.3%
>0.9*HASF	-19.3%	-18.1%	-17.5%	-16.9%	-16.1%	-15.4%	-14.5%	-13.5%
No Dam	-8.3%	-4.2%	-2.0%	0.0%	2.2%	4.4%	6.7%	9.1%

Table 4-4: Median of 2017-2022 percent change in annual average streamflow at Lake Nasser for each filling policy and precipitation trend. All values are relative to the no dam policy and no changes to future precipitation.

<i>filling policy</i>	<i>Precipitation Trend, temp all +2.5 C</i>							
	-20%	-10%	-5%	0	+5%	+10%	+15%	+20%
10%	-9.1%	-7.1%	-6.0%	-4.9%	-3.9%	-2.8%	-1.7%	-0.6%
25%	-15.4%	-13.7%	-12.6%	-11.7%	-10.7%	-9.7%	-8.6%	-7.5%
>HASF	-8.9%	-7.6%	-6.9%	-6.2%	-5.5%	-4.9%	-4.3%	-3.7%
>0.9*HASF	-11.1%	-9.9%	-9.3%	-8.7%	-8.0%	-7.3%	-6.6%	-5.8%
No Dam	-5.8%	-2.9%	-1.4%	0.0%	1.5%	3.0%	4.6%	6.2%

As expected, the percent reduction of annual average streamflow at Gezira Scheme and Lake Nasser over a shorter period (Table 4-3 and 4-4) shows more variability across different filling policies. The most apparent changes occur with the 25% filling policy: under no precipitation trend, the percent reduction increases from 11.5% to 22.8% at the Gezira Scheme and from 5.9% to 11.7% at Lake Nasser. During this filling period, the impounding rate or filling policy, rather than the precipitation change, dominates the downstream flow percent change. For example, at the Gezira Scheme (and Lake Nasser), the difference between the driest climate scenario (-20% precipitation trend) and the no precipitation trend scenario is at most 6.1% (typically less than 5% for Lake Nasser) given the short duration of years; however, comparing policies within each precipitation trend, the gap between the 25% filling policy and no dam scenario increases from 19.5% to 25.9% (from 9.6% to 13.7% for Lake Nasser) as the climate becomes wetter. This is attributed to greater evaporative losses and trends in the other three sub-basins.

Evaluating the full set of simulations for each filling policy and precipitation trend gives a sense of expected variability (Figure 4-8; the Gezira Scheme results not shown but illustrate similar patterns). Interestingly, even at the -5% precipitation trend, each filling policy has some chance of exceeding the historical annual average streamflow. Likewise, for all wetting trends, there is still some probability of reduced flows into Lake Nasser. Additionally, the 25% filling policy still produces the most extreme percent reduction for all simulations across all climate change scenarios. For instance, under no precipitation trend, the percent reduction in annual average streamflow during the filling stage (2017-2032) can be as much as 13.4%.

Note that the model errors assessed from the calibration will carry over accordingly across different filling policies such that the influence of model errors is minimized and the differences

between filling policies are emphasized. We acknowledge that uncertainty in model structure is not explicitly accounted for in this analysis, however, we believe this suffices for the scope of the chapter. Ideally, these results may provide regional decision-makers with a set of plausible outcomes for comparison; however, the sizeable quantity of information (e.g. Figure 4-8) and consideration of many potential future climate conditions may be overwhelming. Clearly, the GCMs tend to favor a slightly wetter future, in general (Figure 4-6). Using this information, weights can be assigned to each precipitation trend, based on a probability density function averaged over the 2056-2065 period providing the likelihood of precipitation changes.

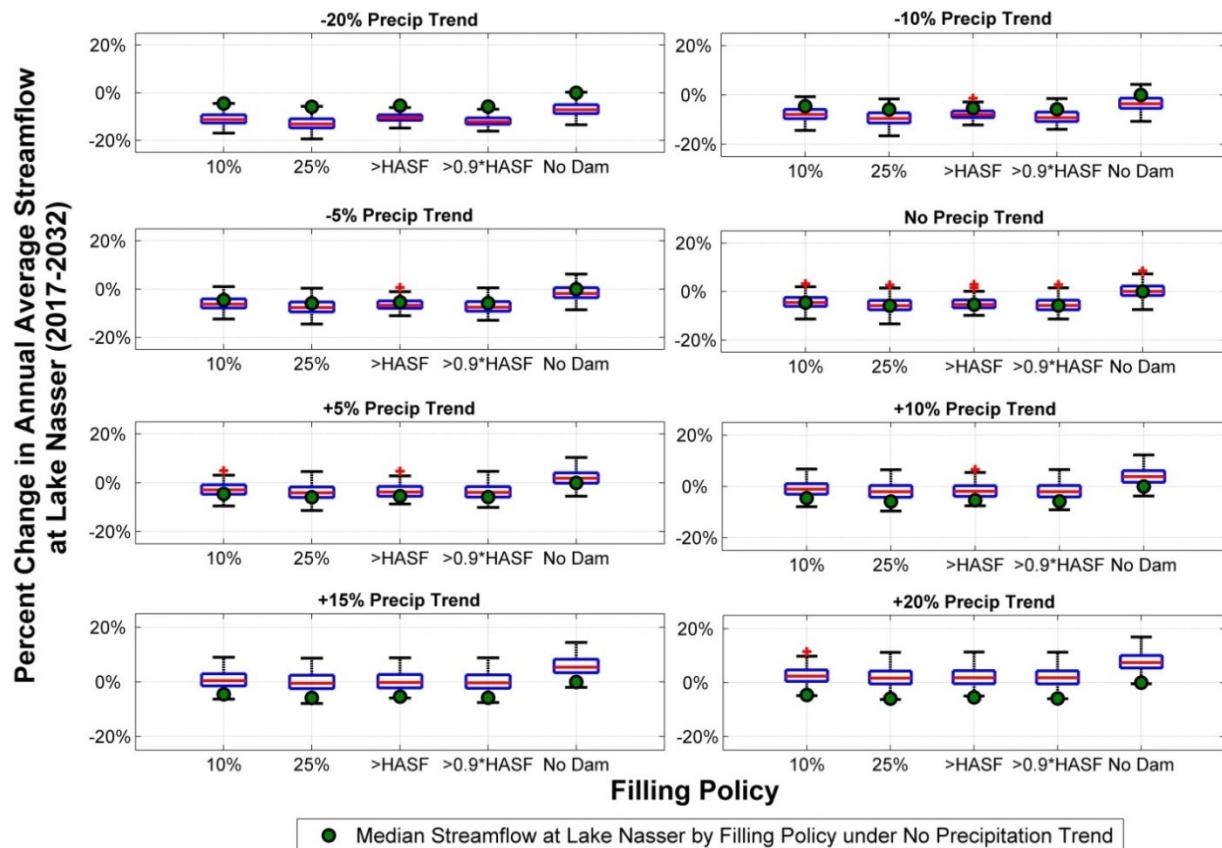
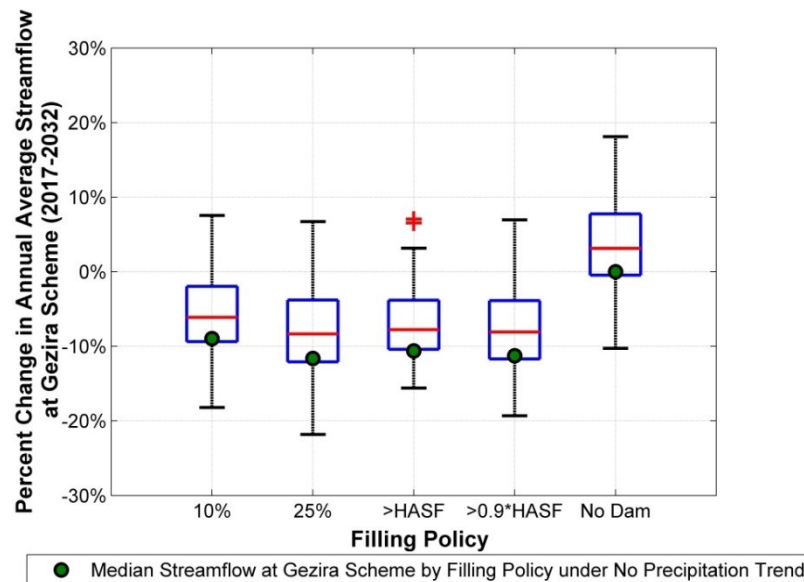


Figure 4-8: Box plots of 2017-2032 percent change in annual average streamflow at Lake Nasser for each filling policy and precipitation trend constructed from the 100 time series. All values are relative to the median no dam policy and no changes to future precipitation. *Note: For each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually in crosses.*

This sensitivity approach effectively collapses the changes in precipitation to an expected value, weighted by the GCM projections (Figure 4-9). In agreement with projected future precipitation changes (Figure 4-6), this produces expected changes in annual average streamflow into the Gezira Scheme and Lake Nasser surpassing the no precipitation trend approximately 75% of the time, for all filling policies.

Also of interest is the significance future changes in precipitation at the end of the simulation period (2060) may have on streamflow reduction (Figure 4-10; Lake Nasser only). All filling policies produce a very similar relationship. Median values under the 25% flow policy are almost indistinguishable from the no dam scenario by 2060, given that the reservoir FSL has been reached after approximately one decade.

(a)



(b)

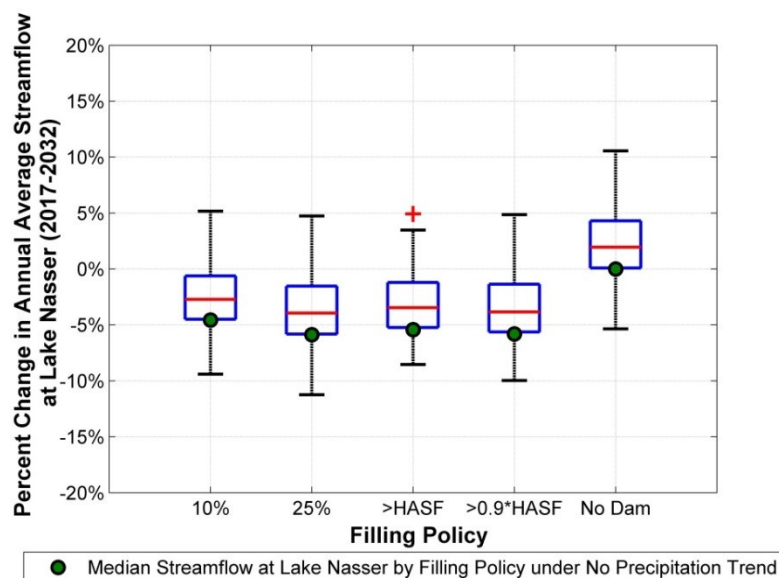


Figure 4-9: Box plots of 2017-2032 percent change in annual average streamflow at (a) Gezira Scheme and (b) Lake Nasser for each filling policy constructed from the 100 time series and weighted precipitation changes. All values are relative to the no dam policy and no changes to future precipitation. *Note: For each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually in crosses.*

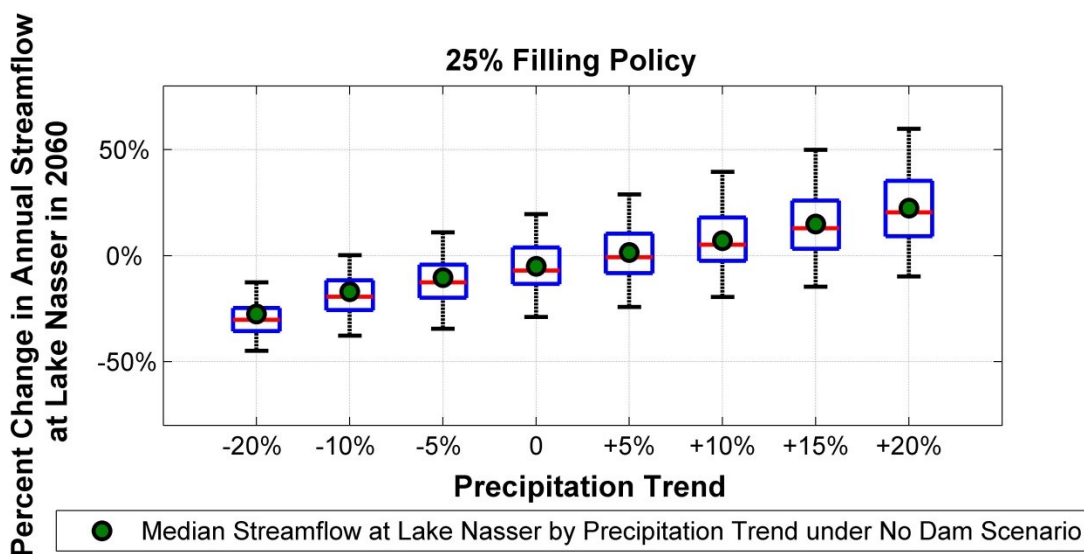


Figure 4-10: Box plots of 2060 percent change in annual average streamflow at Lake Nasser for the 25% filling policy and precipitation trend constructed from the 100 time series. *Note: For each box plot, the short line inside the box is the median, the box edges represent the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually in crosses.*

4.4 Summary and Discussion

The hydrology-hydropower modeling system utilized in this chapter includes streamflow contributions from the Blue Nile, White Nile, Atbara, and Dinder-Rahad to address potential downstream flow impacts given the construction of the GERD. To date, no publicly available filling policy has been announced. Nonetheless, the filling of the GERD reservoir is likely to incur some reduction in streamflow at the Gezira Scheme and Lake Nasser, however the extent varies significantly based on the selected filling policy, climate variability, and projected climate changes during the filling stage. The filling policy, net evaporation, and climate variability dominate reductions in the early years and decades prior to reaching the FSL (Table 4-3 and 4-4); climate change and net evaporation dominate once the FSL has been attained (Table 4-1 and 4-2). How the GERD is filled and operated – whether it is through mutual cooperation or through isolated determination – will set a precedent for the many forthcoming projects across the Nile Basin.

A climate sensitivity approach is adopted in this chapter in lieu of traditional GCM downscaling. The motivation for this comes from a desire to assign probabilities to streamflow reduction for each filling policy, providing decision-makers with a set of plausible, justifiable, and comparable outcomes for negotiation and consensus. This approach lessens the reliance on any single GCM output, especially important given the challenges of GCMs in producing consistent precipitation projections.

The operational strategy of the GERD may also change the annual hydrologic cycle, particularly for downstream countries. This may be advantageous, as downstream flows will become more

regularized month-to-month, potentially facilitating irrigated agriculture and hydropower generation. The GERD will also retain significant amounts of silt within the reservoir, likely extending the effective life of downstream hydropower facilities. Overall, it increases the opportunity for mutual benefits in the Nile Basin and sets a precedent for future basin projects.

Future phases of this chapter include extending the model to demonstrate more specific downstream impacts, such as the effect of the filling rate on the storage in Lake Nasser, associated changes in water allocation at the Gezira Irrigation Scheme, optimal operational strategies for a given filling policy, and a cost-benefit analysis from the perspectives of both Ethiopia and downstream riparian countries—Sudan and Egypt. Joint analysis of GERD and Aswan Dam may be worthwhile, as the impact of filling the GERD on streamflow reduction at Lake Nasser can be alleviated given that the Aswan Dam can potentially store up to approximately twice of the historical average inflow prior to the filling of GERD. Additionally, broader studies will include other planned water infrastructure and developments in the basin, including irrigated agriculture, hydropower generation, and domestic/industrial demand to map benefits and impacts.

Chapter 5 Filling the GERD: Evaluating Hydroclimatic Variability and Impoundment Strategies for Blue Nile Riparian Countries

5.1 Introduction

Talks between riparian countries of Blue Nile River have been on-again, off-again since 2011, shortly after construction of GERD commenced, ranging from handshakes to military threats (Maher, 2013; Gebreluel, 2014; Salman, 2016). Ethiopia has incentive to fill the reservoir rapidly to begin generating hydropower, while downstream countries prefer a modest filling rate to minimize the possible impacts on dependent lives and livelihoods. Although filling is likely to begin as early as 2018 (one year postponed from original plan), no consensus on an agreeable filling strategy has yet been reached. Deliberations are being closely watched by other countries in the Nile Basin and beyond, as cooperation and coordinated action, or lack thereof, regarding filling the reservoir and eventual operation of the GERD is likely to set a precedent for future development in the basin.

On the surface, reservoir filling is a relatively simple matter of water balance. The two key considerations are inflow to the reservoir, based on hydroclimatic conditions, and outflow, based on impoundment/release strategies and, to a lesser extent, evaporation. This chapter addresses both aspects across the reservoir filling stage. Complexities arise when allocations and externalities are considered, both temporally and spatially, specifically who gets water when. To better understand and compare the diverse effects of climate variability coupled with filling strategies, we evaluate Ethiopian hydropower generation at the GERD balanced against storage and downstream releases for a variety of scenarios. In Chapter 4, streamflow reductions were

additionally determined for Sudan's immense irrigation scheme at Gezira and Egypt's Lake Nasser; in this analysis, downstream flow conditions are restricted to GERD releases, from which streamflow into Gezira and Egypt can be inferred (Chapter 4). Adaptive operational strategies in Sudan and Egypt in coordination with the GERD release strategies may improve in-country water management; however, this coordinated operation is not explicitly addressed in this chapter; the reader is referred to Wheeler et al. (2016) for details.

In addition to contributing to the larger, ongoing discussion surrounding the GERD, this chapter is unique on two fronts. First, it accounts for both interannual and low frequency variability in hydroclimatic variables in construction of projected streamflow across the filling stage; and second, it considers numerous filling strategies. Streamflow simulation methodologies, filling strategy specifics, model structure, results, and a discussion are provided in the following sections.

5.2 What Does the Future Hold? Projecting Hydroclimatic Conditions

Precipitation across the Blue Nile basin and streamflow at the GERD site has strong intra- and interannual variability. The annual precipitation cycle is dominated by the Kiremt (main) rainy season, spanning June-September, which contributes approximately 70% of annual basin-wide precipitation on average (Figure 5-1a). While average annual total precipitation is nearly 1270 mm, year-to-year variability is notable, with a coefficient of variation approximately 12% (Figure 5-1b). Historically dry years (e.g. 1984 and 2002) are often associated with El Niño conditions, occurring at periodicities in the 3-8 year band (e.g. Bekele, 1997; Wolde-Georgis, 1997; Block and Rajagopalan, 2007; Diro et al., 2011a). Multi-decadal variability resulting from

lower frequency oscillations can produce extended periods of above or below normal precipitation; however, evaluation of Blue Nile precipitation at this scale has received relatively less attention, given that the low frequency signals explain a smaller fraction of the total variance. When projecting expected reservoir inflow conditions (e.g. high or low) across the filling phase, however, it may be important to capture significant signals at all frequencies given the potential implications for filling rates and storage, downstream flows, hydropower generation, etc. Anomalously high flow conditions across filling years could produce drastically different outcomes than if low flow conditions persist, for example. Thus investigation into the presence of low frequency signals and subsequently projecting inflow conditioned on interannual and low frequency signals is warranted.

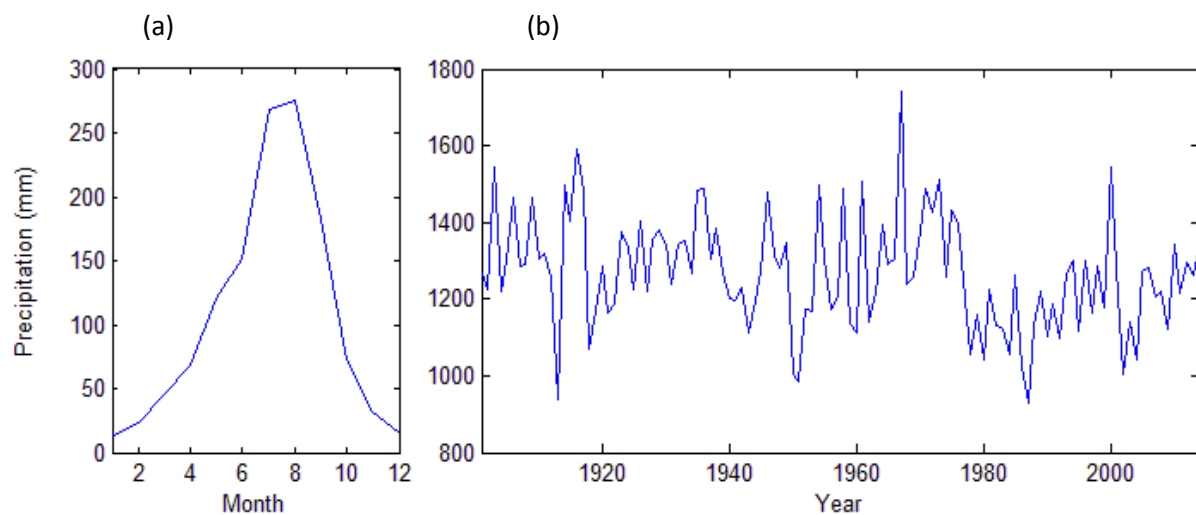


Figure 5-1: Blue Nile basin precipitation: (a) monthly average, and (b) annual total based on CRU TS v3.23 data set (<https://crudata.uea.ac.uk/cru/data/hrg/>)

Wavelet analysis is selected to evaluate the existence of low frequency signals in average, basin-wide precipitation over the Blue Nile Basin using the CRU TS v3.23 gridded global monthly precipitation dataset from 1901 to 2014 (Harris et al., 2014). Wavelet analysis provides a mechanism to investigate precipitation in time and frequency domains simultaneously, such that

significant signals at certain frequencies and times can be identified. A Morlet mother wavelet is chosen for its properties suitable for precipitation time series analysis (Kwon et al., 2007). For a more detailed description of wavelet analysis, see Kumar and Foufoula-Georgiou (1997) and Torrence and Compo (1998). For data preparation, the precipitation time series is aggregated into annual values to reduce seasonality effects and potentially emphasize any low frequency signals. It is then detrended and normalized for wavelet analysis. In this chapter, wavelet analysis produces two low frequency bands that are significantly different from the background noise spectrum (i.e. white noise), specifically on 21-year (90% significance) and at 35-year (95% significance) periods, as indicated by the peaks in the wavelet spectrum (Figure 5-2a, b).

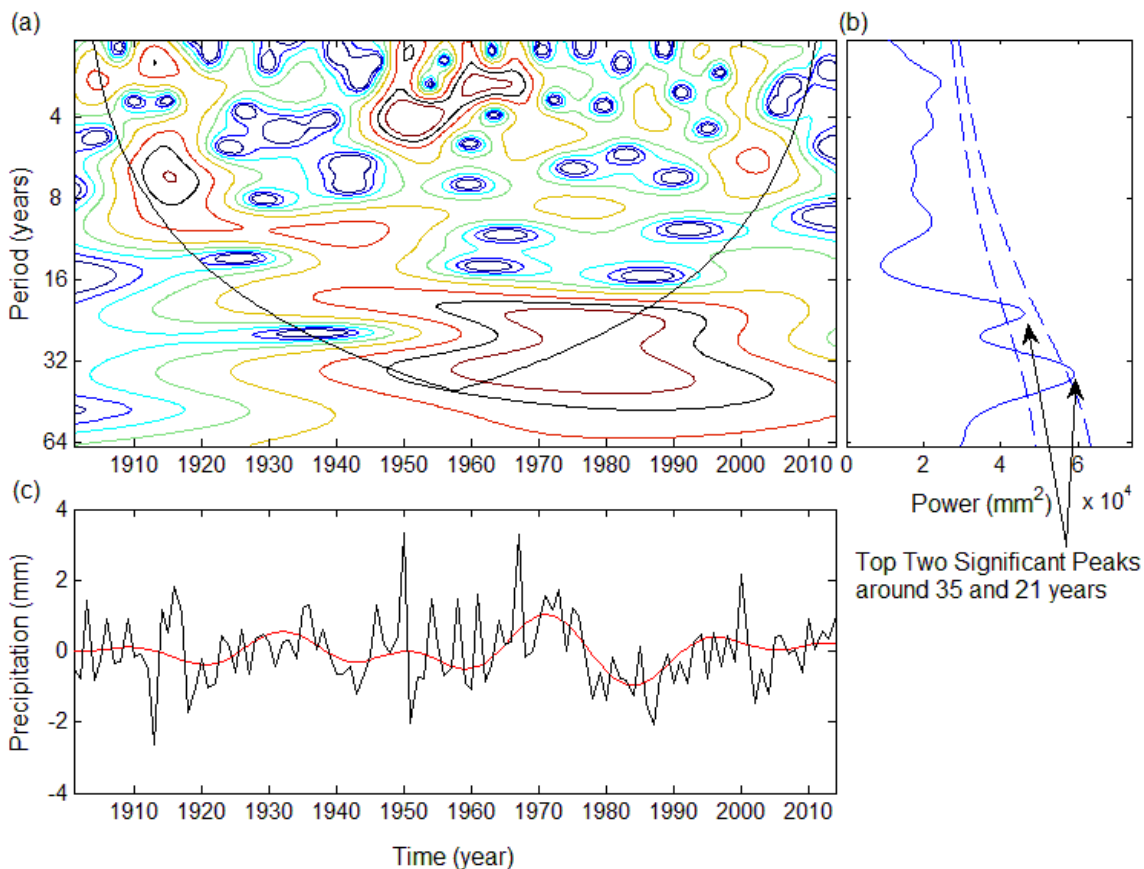


Figure 5-2: Wavelet analysis of Blue Nile annual precipitation: (a) wavelet power spectrum, (b) global wavelet spectrum, with dashed lines showing the 90% (left) and 95% (right) significance levels and (c) standardized detrended annual precipitation (black solid line), with combined low-frequency signal reconstructed at 21-year and 35-year periodicity, scaled.

Given the presence of both interannual and low frequency variability, inflow projections across the filling period are conditioned on both, as described below. Projections start in 2015, the end of the precipitation dataset, and continue to 2034. Filling of the GERD reservoir, however, is not expected to commence until 2018. Accordingly, 204 months (2018-2034) is the full period allowable for filling, while the full simulation period is 240 months (2015-2034). Beginning with the low frequency periods, time series ($S21$ and $S35$) associated with each significant peak can be reconstructed (Torrence and Compo, 1998) and then scaled for the historical period (e.g. Figure 5-2c):

$$X_t = \alpha_1 S21_t + \alpha_2 S35_t + \varepsilon_t \dots\dots\dots \text{(Equation 5-1)}$$

where X is the detrended, normalized annual time series, t is the yearly time index, α_1 and α_2 are fitted coefficients (i.e. the scaled factors for each reconstructed signal), and ε is the residuals from linear regression.

The scaled $S21$ and $S35$ time series are projected beyond 2015 independently using a non-parametric Block-KNN method (Erkyihun, 2015), and subsequently summed. The Block-KNN projection requires the following steps, illustrated in Figure 5-3 for the 21-year low frequency signal:

- 1) Define the phase length P of one oscillation in the signal, equal to 21 or 35 years in this case. P is determined from the global wavelet power spectrum (e.g. Figure 5-2b)
- 2) Using the final $P/4$ length block (Erkyihun, 2015) at the end of the historical time series, find the closest matching block with the same length from the previous years using a k -nearest neighbor (KNN) bootstrap technique (Lall and Sharma, 1996).

- 3) From the end of the matching block, select the *next* 20-year length block (projection) and append it to the end of the historical times series, creating a projection to 2034. The projected block should be shifted vertically (correcting any bias) to create a smooth transition with the reference block.
- 4) Repeat steps 1 - 3 for each significant frequency and then sum the projected time series to form an aggregate annual low frequency projection.

For this chapter, a single projection is proposed for each of the two significant low frequency signals identified, given the limited amount of variance they explain (17%). Alternatively the uncertainty associated with the low frequency projections could be accounted for by repeating the Block-KNN method many times and selecting from the full range of nearest neighbors (not simply the closest) to create a projection envelope.

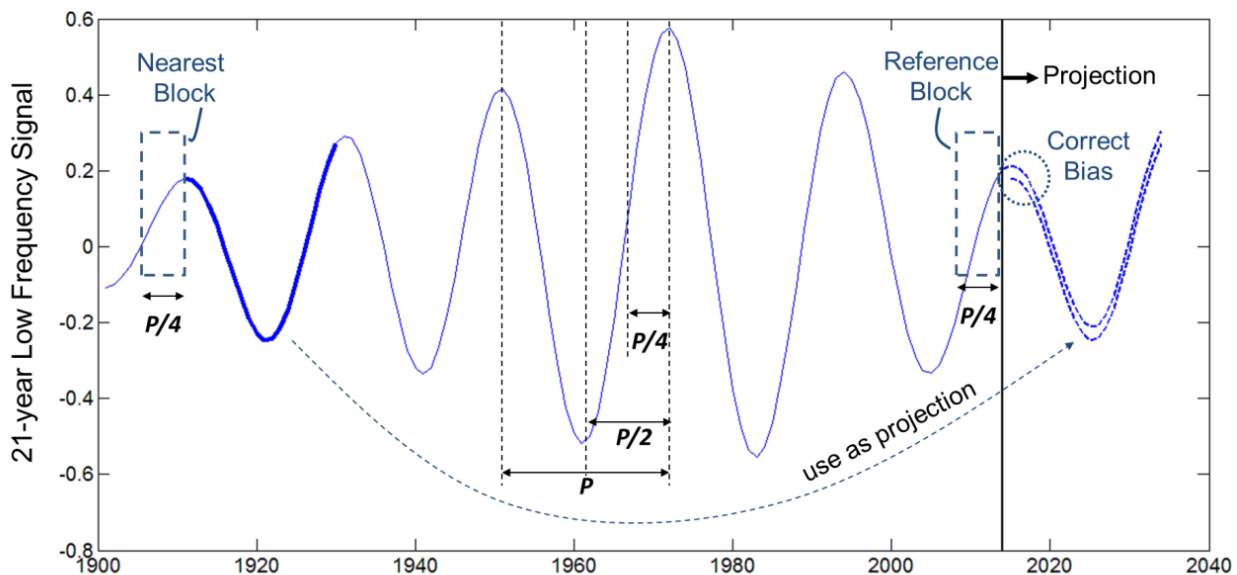


Figure 5-3: Illustration of the Block-KNN projection methodology.

In contrast, to better represent the possible range of precipitation within each month, the intra-annual precipitation data are projected stochastically. This requires a monthly time series with the low frequency signal removed (Equation 5-2). Thus, the annual precipitation time series X is disaggregated into a monthly time series and then the low frequency contribution, $S_t/12$, is removed. Note that precipitation values associated with the low frequency signal are assumed to be equivalently distributed across all months in a year; that is, seasonality is preserved in the residual monthly time series.

$$M_{i,t} = \frac{Y_{i,t}}{\sum_{i=1}^{12} Y_{i,t}} X_t - \frac{S_t}{12} \dots\dots\dots \text{(Equation 5-2)}$$

where $M_{i,t}$ is the residual monthly time series at month i in year t and $Y_{i,t}$ is the original monthly precipitation time series. Figure 5-4 illustrates the aggregated low frequency signals and annual and monthly residuals across the historical record.

To construct 20-year projections (2015-2034), stochastic simulations (100) of interannual variability are added to the static low frequency projection. For each simulation, 20 years are randomly bootstrapped with replacement from the historical period (1901-2014); for each year, the 12 monthly values, $M_{i,t}$, are added to the corresponding (year t) aggregated low frequency value, $S_t/12$, to form monthly time series with the low frequency signal embedded (Figure 5-5 illustrates one simulation).

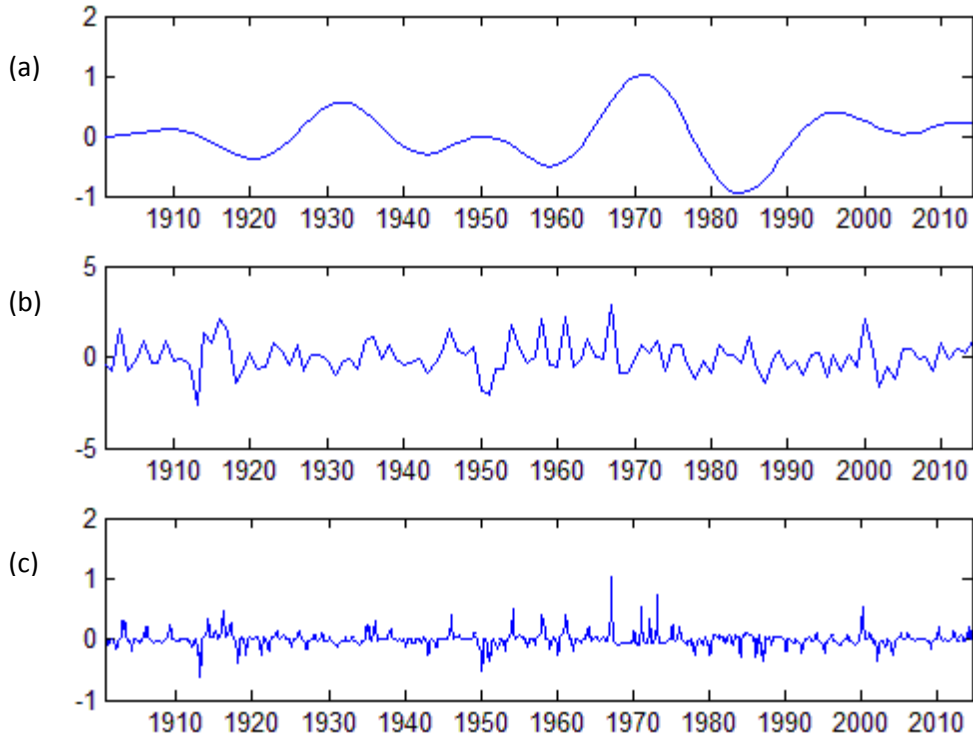


Figure 5-4: (a) Aggregated low frequency signal in Blue Nile annual precipitation over 1901-2014 ($\alpha_1 S21_t + \alpha_2 S35_t$; see Equation 5-1); values are scaled. (b) Residual annual time series after subtracting the signal in (a) from the detrended normalized historical record. (c) Residual monthly time series after disaggregating from annual time series to monthly time series and subtracting the signal in (a) using Equation 5-2.

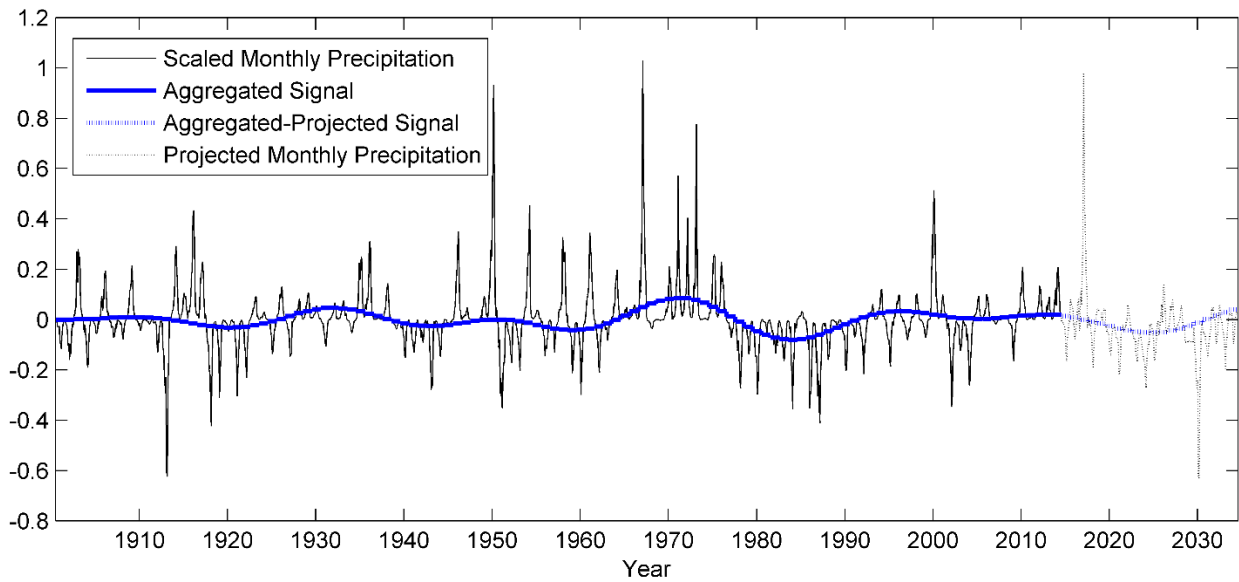


Figure 5-5: Scaled monthly precipitation from the historical and projected periods, highlighting the embedded low frequency signal (single simulation only).

Corresponding monthly temperature time series projections are also needed in addition to the precipitation projections, to eventually generate streamflow. To preserve the relationship between precipitation and temperature, one additional step is required. For each month of the precipitation projections, the month with the most similar precipitation value (in the same month, e.g. Januaries only) from the historical (CRU TSv3.32) record is selected and both precipitation and temperature from that month are retained. These, then, form the final set of projections.

5.3 How Will the Reservoir be Filled? Possible Filling Strategies

The rate at which the reservoir is filled has direct implications for water storage, downstream releases, and hydropower generation. Three types of filling strategies are considered here, including fractional, threshold, and absolute (absolute is new from Chapter 4). These strategies, further described below, are not unique to this chapter, but simultaneous comparison across these three categories has not been undertaken.

Fractional filling strategies allow impounding a specified percent of total monthly streamflow into the GERD reservoir; both 10% and 25% are evaluated here. This strategy guarantees that some quantity of water can be impounded, although this quantity varies month to month and year to year (King and Block, 2014). Fractional strategies generally favor a sharing by upstream and downstream countries of the risk associated with streamflow variability. In this case, the general water balance is:

$$V_t = V_{t-1} + X\% * Q_t - E_t \dots\dots\dots \text{(Equation 5-3)}$$

where V is reservoir volume, t is time (month), $X\%$ is the fraction impounded (here, either 10% or 25%), Q is streamflow volume into the reservoir, and E is net evaporation volume from the

reservoir.

The threshold filling strategy allows any streamflow volume in excess of the long-term historical monthly average to be impounded in the reservoir. Thus, in months with anomalously high flow, the volume impounded can be large; however, in months with flow below the average, no impounding is permitted (King and Block, 2014). In months with flow below the average, Ethiopia is not required to make up the difference with existing reservoir storage. This strategy generally favors downstream countries, as water for impoundment in any given month is not guaranteed, and downstream countries receive at least long-term average monthly streamflow or the full flow volume in drier than average months. In this case, the general water balance is:

$$V_t = V_{t-1} + \max(Q_t - HASF_t, 0) - E_t \dots\dots\dots \text{(Equation 5-4)}$$

where *HASF* is the long-term historical average streamflow volume into the reservoir.

Absolute filling strategies allow for a guaranteed volume of water to be impounded in the reservoir annually. These strategies are structured here based on time taken to fill the reservoir, namely 4, 6, or 8 years. To fill the reservoir in 4 years, for example, one-fourth of the total reservoir volume may be impounded each year, irrespective of the streamflow. Annual flow impounded is disaggregated to monthly flow weighted by the long-term historical monthly average, and then adjusted by a factor γ such that the expected time to fill (e.g. 4, 6, or 8 years) can be guaranteed (minus the effect of evaporative loss). It is possible, therefore, that in abnormally low flow months, insufficient water will be available to meet the full impoundment demand, and all water will be stored with no releases downstream. These strategies generally favor upstream countries, as water for impoundment in any given month is guaranteed and the expected time to fill the reservoir is fixed. In this case, the general water balance is:

$$V_t = V_{t-1} + \gamma * U_t - E_t \dots\dots\dots \text{(Equation 5-5)}$$

where U is the weighted monthly impoundment volume and t is in months.

Thus, six filling strategies (of three general types) are tested. It is important to note that all water not impounded may be passed through turbines to generate hydropower.

5.4 From Climate to Storage to Allocation: Model Structure

To translate the temperature and precipitation time series into monthly streamflow entering the GERD reservoir, a variation of the WatBal water balance model (Yates, 1996) is adopted as calibrated in Chapter 4 using the same dataset from CRU (Harris et al., 2014). WatBal was originally developed to model Nile River Basin response to climate change impacts and has been shown to satisfactorily reproduce monthly streamflow observations along the Nile River (Yates and Strzepek, 1998; Chapter 4). WatBal includes three modules: a soil moisture modeling scheme to account for water fluxes, potential evapotranspiration using the Hargreaves method (Hargreaves and Samani, 1982), and a storage scheme for lakes and swamps. Additional details are given by Yates (1996) and Chapter 4.

The GERD dam and reservoir model simulates monthly storage volume, hydropower generation, and releases for downstream countries contingent on the filling strategy selected. The time to fill the reservoir is also computed, and can serve as a useful metric in comparing various strategies and the effects of climate variability. Although the model is run for 20 years, representing 2015-2034, normal operating rules likely to be implemented beyond the filling period are not included. Dam and reservoir design characteristics were developed from publicly available sources, the

International Panel of Experts report (Elsayed et al., 2013) and a preliminary study conducted by the United States Bureau of Reclamation (USBR, 1964). Original plans allow for up to 9 km³ of storage by 2015 with two 375 MW turbines online to generate hydropower (EEPCo, 2013). Upon dam completion in 2018, further storage may continue up to the full supply level of 74 km³, and all turbines will be available for generating hydropower, up to a final capacity of 6,000 MW.

5.5 Results

5.5.1 Time to Fill the Reservoir

The time required to fill the reservoir to the full supply level, while less critical from an operational perspective, is a tangible metric of interest to policy makers in the countries. The starting point is January 2018 – with an initial volume of 9 km³ as previously discussed – and the ending point (when the full supply level is initially reached) is contingent on the filling strategy. It is worthwhile to note that after the reservoir is filled, the filling strategy adopted will no longer be practical for operations; that is, a different operational strategy will be adopted. However, given the focus here on the filling stage, operational strategies after filling are not explored. Thus while all model runs cover the same projection period (2015 – 2034, 240 months) consistently for each filling strategy, the results (reservoir volumes, downstream flows, and hydropower generation) after the reservoir is filled may not be realistic; hence, they are truncated from the figures for months after the reservoir is filled.

Based on the 100 simulations with the low frequency signal embedded, the 4-year filling strategy, not surprisingly, tends to fill the most quickly, with a median filling time of only 47

months and 95% of the simulations filling within 54 months. The median time to fill is approximately equivalent for the 25% (68 months), HASF (68 months), and 6-year (71 months) filling strategies (Figure 5-6, Table 5-1). However, considering the time to fill associated with the lower 95th percentile based on the simulations (Table 5-1), while the 25% and 6-year strategies differ by only 8-12 months from their respective medians, the HASF strategy requires more than 204 months, implying that the reservoir would not fill even by 2034. Thus the 25% and 6-year strategies have relatively low variance in their time to fill, which is advantageous for reservoir planning; in contrast, the HASF strategy has a large variance in time to fill, under which downstream streamflow would remain relatively close to the historical average each month. The 10% strategy clearly requires the longest time to fill the reservoir.

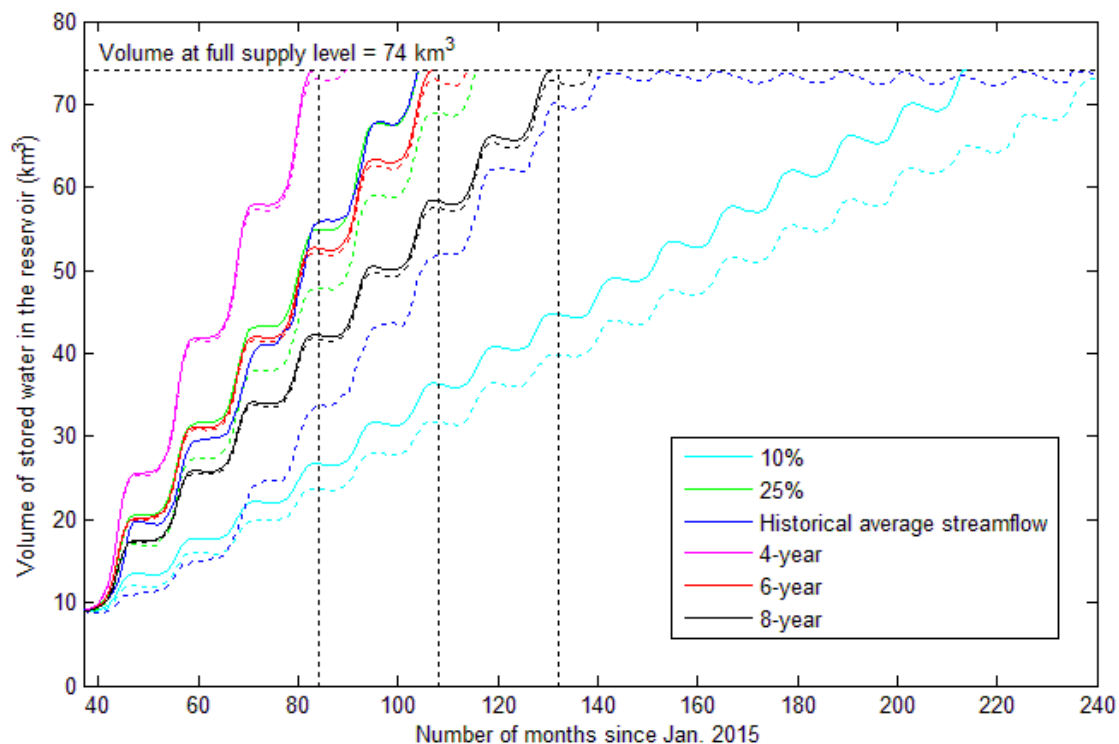


Figure 5-6: Volume of water in the reservoir over time for different filling strategies based on projections with the low frequency signal embedded. Solid and dashed lines indicate the median and lower 95th percentile, respectively, based on the 100 simulations.

Table 5-1: Time to fill the reservoir for different filling strategies.

Filling strategy \ Time to fill		10%	25%	HASF	4-yr	6-yr	8-yr
Projection with Low Frequency Signal Embedded							
Number of months to fill	Median	178	68	68	47	71	95
	Lower 95 th percentile	>204	80	>204	54	79	103
	<i>Median-95th</i>	>26	12	>136	7	8	8
Median year filled		2032	2023	2023	2021	2023	2025
Projection without Low Frequency Signal Embedded							
Number of months to fill	Median	188	68	67	47	71	95
	Lower 95 th percentile	201	80	140	54	79	103
	<i>Median-95th</i>	13	12	73	7	8	8
Median year filled		2033	2023	2023	2021	2023	2025

The effect of explicitly including the low frequency signal on the time to fill is also of interest, particularly compared to projections without the low frequency signal included, produced here by randomly bootstrapping years from the historical record to construct the 20-year simulations. Comparing the two approaches, the median time to fill is relatively similar for each filling strategy; however, excluding the low frequency signal results in a vastly narrower range of uncertainty (smaller spread between time to fill outcomes for the 100 simulations) for the 10% and HASF strategies. When the low frequency signal is ignored, these two strategies can fill the reservoir before 2034 even at the 95th percentile, in sharp contrast to the projections accounting for the low frequency signal. Even at the median time to fill, the differences for these two approaches are evident (Table 5-1). Although the time to fill is the same with and without the low frequency signal for the 25% and absolute filling strategies, the monthly storage and release volumes still differ (results not shown). Therefore if a filling strategy is selected that is based more heavily on current streamflow conditions – and not preset or absolute withdrawals – accounting for the low frequency patterns in streamflow becomes increasingly important.

5.5.2 Downstream Releases during the Filling Stage

Mean annual downstream releases from the GERD vary between 33 – 46 km³ at the beginning of the filling stage, dependent on the filling strategy, when the low frequency signal is embedded (Figure 5-7). The 4-year strategy clearly has the largest initial abstractions, resulting in significantly reduced flows downstream; however, flows return to normal faster than under any other strategies. In contrast, the 10% strategy produces relatively stable releases fluctuating around 45 km³, not returning to normal flow until the end of the projection period. Considering the lower 95th percentile of simulations, releases fall below 15 km³ under the 4-year strategy, and even below 30 km³ under the 10% and HASF filling strategies. Here the HASF strategy produces the largest releases throughout the filling stage, on average, and also represents the smallest gap relative to median releases. This relatively low uncertainty is in contrast to the large HASF strategy uncertainty in time to fill, and emphasizes its favorableness to downstream countries.

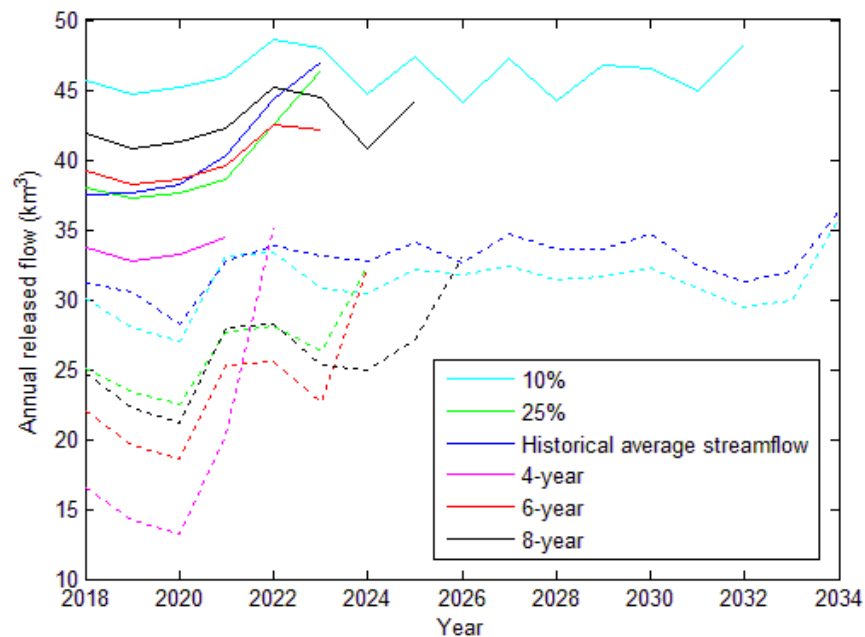


Figure 5-7: Annual downstream releases from the dam for different filling strategies based on projections with the low frequency signal embedded. Solid and dashed lines indicate the mean and lower 95th percentile, respectively, based on the 100 simulations.

The percentage reduction in annual downstream releases for each filling strategy (each simulation compared with historical averages) illustrates the variance expected each year throughout the filling stage (Figure 5-8). Unsurprisingly, for the percentage-based filling strategies, the percentage reduction during the filling stage is simply the percentage itself (10% and 25%) for each simulation resulting in no variability. For the absolute filling strategies, the variability of relative reduction is higher under the 4-year strategy than under the 6 or 8-year strategy. Though the HASF strategy has the highest variability among all filling strategies during the filling stage, the absolute amount of streamflow downstream of the GERD is often simply the long-term historical average; for this strategy, reductions imply that water was impounded during a wet year. Finally, the median reductions for the 4-, 6-, and 8-year strategies are approximately 35%, 24%, and 18%, respectively. The consistency between time to fill and downstream releases is evident by comparing Figures 5-6 and 5-8.

The multi-year shortfalls in GERD releases are also a key concern for downstream countries. Correspondingly, minimum releases for continuous 12, 24 and 36-month (1, 2 and 3-year) spans under each filling strategy are computed, and the lower 95th percentile and median values from the 100 simulations are presented (Table 5-2). The results indicate the possible severity of downstream reductions across the filling strategies. Possible reductions are clearly largest for the absolute filling strategies; the fractional-based and HASF strategies, fluctuating with available streamflow, produce lesser possible reduction.

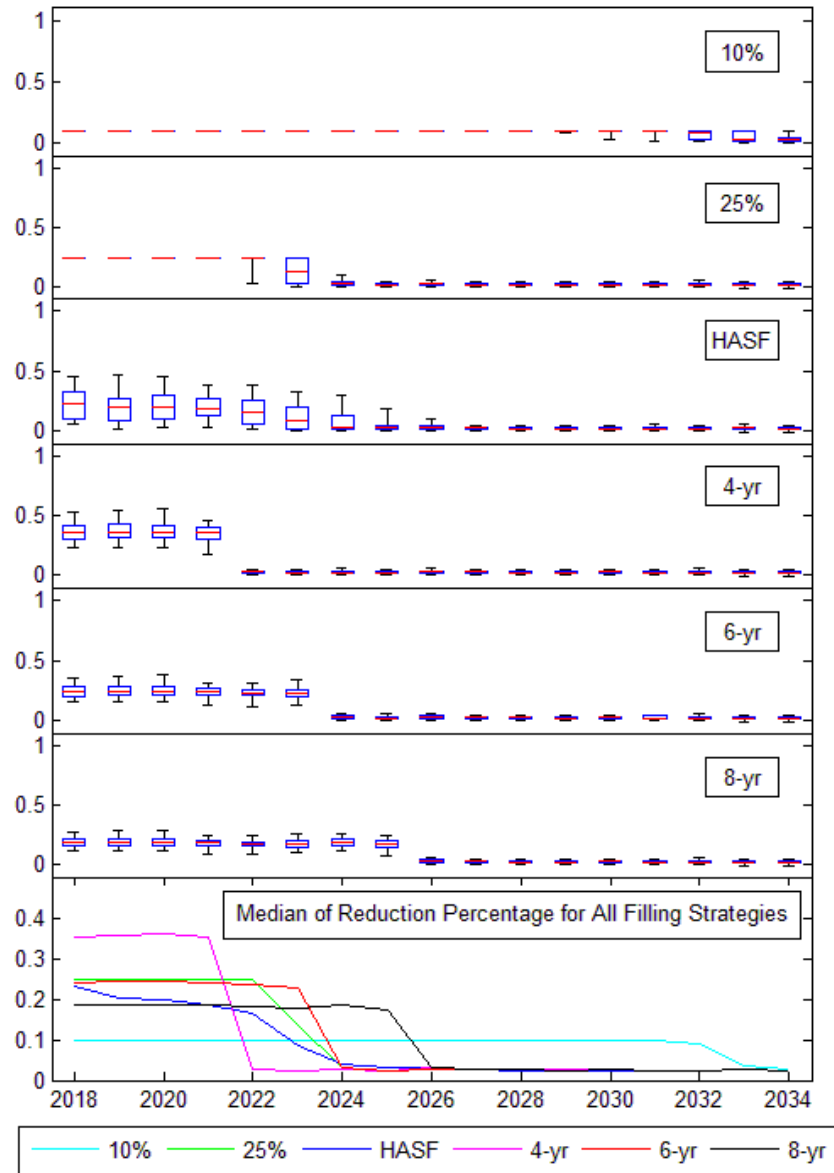


Figure 5-8: Percentage reduction in annual downstream releases for each filling strategy, compared with historical averages; the low frequency signal is embedded. Boxplots are based on 100 simulations. *Note: For each box plot, the line inside the box is the median, the box edges represent the 25th and 75th percentiles, and the whiskers extend to the 5th and 95th percentiles.*

Table 5-2: lower 95th percentile and median values of annual GERD releases (1, 2 and 3-year shortfalls in km³) for each filling strategy; 24 months and 36 months are per year averages.

Multi-yr Shortfalls	Lower 95th percentile Median											
	10%		25%		HASF		4-yr		6-yr		8-yr	
12 months	23.1	29.1	19.6	26.3	23.3	30.6	10.6	20.0	14.6	24.3	17.3	25.1
24 months	27.9	33.3	24.9	30.8	29.5	34.4	16.3	25.2	21.8	29.3	23.7	30.3
36 months	30.6	36.1	28.0	32.9	31.7	35.7	20.4	28.7	25.9	32.1	26.8	33.4

As previously mentioned, under the absolute filling strategies months with no downstream releases are possible. This occurs during an exceptionally low flow month, when streamflow is less than the allotted impoundment volume. Not surprisingly, these months occur more often under the 4-year strategy than under the 6- or 8-year strategy. The no release months occur most frequently in February and March, the dry season in the Blue Nile basin (Figure 5-9). For example, only nine no flow months occur under the 8-year strategy, summing across all 100 simulations; while this is minimal, it is by no means trivial. Shortening the filling stage increases the number of no flow months and distributes across other months; under the 4-year strategy, some no flow months occur even during the high flow season (September-October) in a few simulations. Thus while the absolute filling strategies have small uncertainty in the time to fill, the possibility of months with no downstream flow may pose a serious threat to downstream countries.

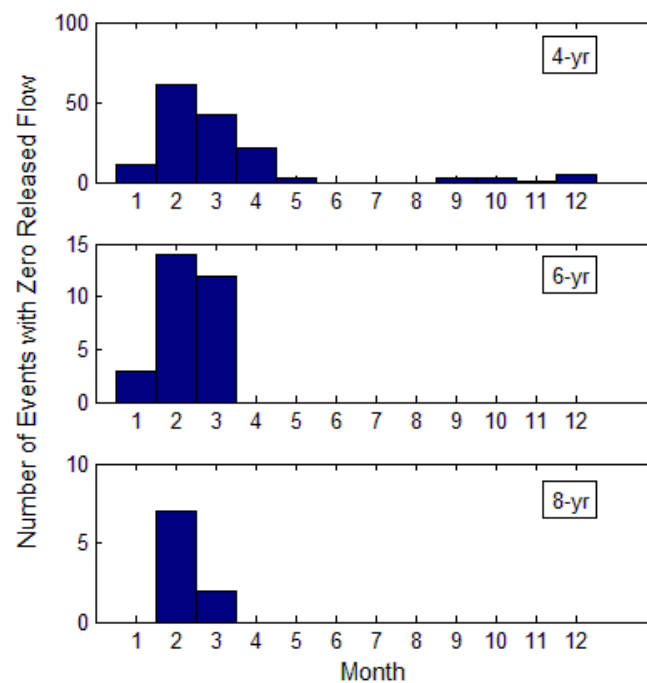


Figure 5-9: Number of months with no downstream releases for the three absolute filling strategies based on projections with the low frequency signal embedded, accumulated from the 100 simulations. *Note: the y-axis scale differences for each subplot.*

The number of no flow months under each absolute strategy for each projection, both with and without the low frequency signal included, can be quite different (Figure 5-10). When no low frequency signal is embedded in the projections, no flow months occur much more frequently. For the 4-year strategy, the maximum number of months with no flow varies between 1 and 16 for projections without the low frequency signal; at the high end, this is an average of four months per year across the filling stage. When the low frequency signal is included, there are a maximum of six no-flow months. Even though the 8-year strategy with the low frequency signal included has a relatively low number of simulations with no flow months (6%), these cannot be ignored; that is, one cannot claim at the 95% confidence level that zero no flow months would occur. Under the conditions assumed for this chapter, projections including the low frequency signal are generally associated with larger streamflow, compared to projections without the low frequency signal included, particularly during the dry months (January, February and March), resulting in fewer no flow months. Essentially, including the low frequency signal produces time series with more moderate years, that is, fewer extremely low-flow and high-flow months. Therefore, there are fewer months having zero releases, due to the smaller number of extremely low flow months; yet, more time is required to fill the reservoir since the number of high flow months is also smaller.

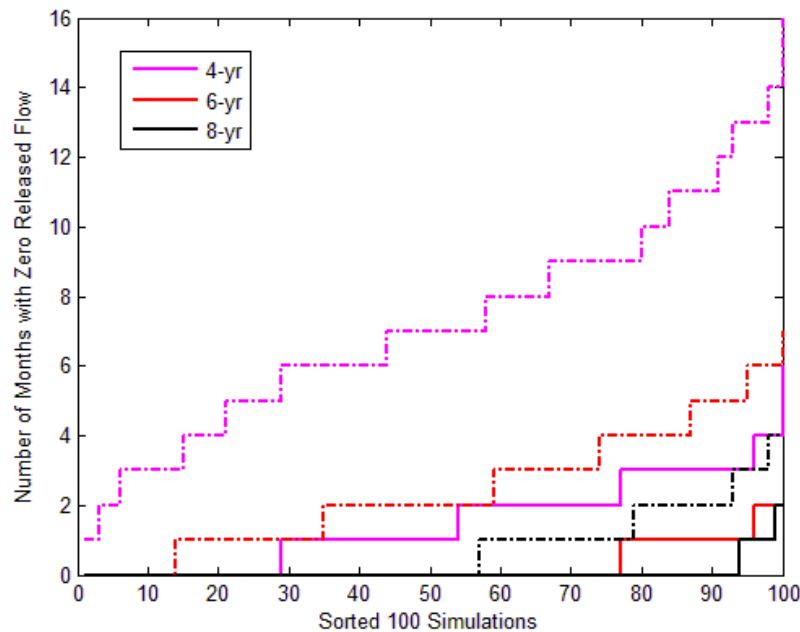


Figure 5-10: Number of months with no flow released downstream of the dam for three absolute filling strategies, accumulated for all months in each simulation (ordered). Solid and dash-dot lines indicate the results based on projections with and without the low frequency signal embedded, respectively.

5.5.3 Hydroelectric Generation

Given the 9 km^3 allowable reservoir storage volume prior to the start of the official filling stage (January 2018 in this chapter), only an additional 5.8 km^3 of storage is required to reach the minimum operation level from which hydropower generation may commence at full speed (Elsayed et al., 2013). For most filling strategies, this occurs in the first year of filling; annual hydroelectric generation then increases rapidly through the filling stage (Figure 5-11). Annual hydroelectric generation varies from 9,000 GWh to 10,400 GWh, with strategies releasing more water (storing less) typically generating more energy (Table 5-3). All filling policies have simulations in which they generate 11,000 GWh or more, averaged over the filling stage, with the exception of the 4-year strategy, which has a maximum of 10,000 GWh across the filling

stage (Table 5-3). This is not surprising, as the 4-year strategy emphasizes storing water and releasing less downstream flows. However, because the 4-year strategy fills the most quickly, the ensuing years are likely to produce more hydropower than under other strategies. For example, the 13,600 GWh produced, on average, under the 4-year strategy across 2018-2034 surpasses all other strategies. (This result should be used primarily for comparison with other strategies and not in terms of absolute values, given the aforementioned discussion regarding operational policies after filling. For the calculations here, once the reservoir is filled, generation is expected to be approximately 15,000 GWh/year (Elsayed et al., 2013).)

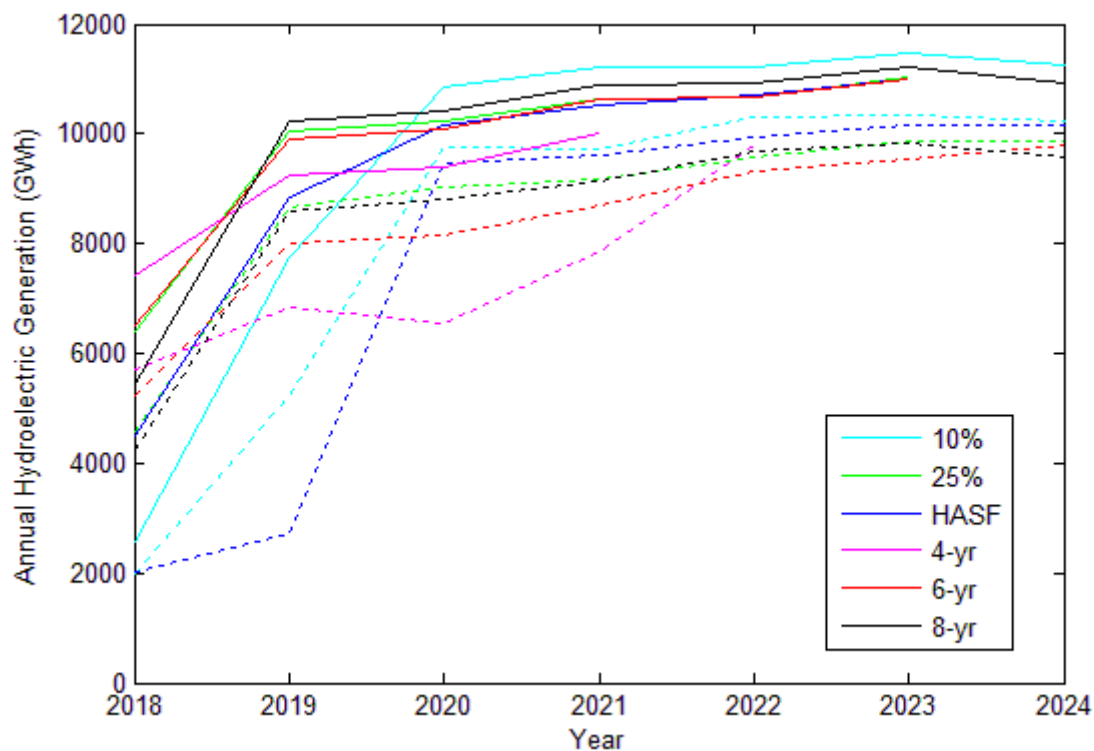


Figure 5-11: Annual hydroelectric generation based on projections with the low frequency signal embedded. Solid and dashed lines indicate the mean and lower 95th percentile of 100 simulations, respectively.

Table 5-3: Annual hydroelectric generation (GW-hrs) during the filling stage (unique to each strategy) and full projection period (2018 - 2034), assuming full hydroelectric generation after the reservoir is filled (15 000 GW-hrs annually). Calculations are based on the mean values of 100 simulations and numbers are rounded to hundreds.

Filling Strategy		10%	25%	HASF	4-yr	6-yr	8-yr	
Hydroelectricity	during filling stage	mean	10 400	9 800	9 300	9 000	9 800	10 100
		max	11 500	11 000	11 000	10 000	11 000	11 200
2018 - 2034		mean	11 000	13 200	13 000	13 600	13 200	12 700

It is also important to consider the discounting effect on hydroelectricity benefits; i.e. the hydropower generated early will be more valuable than that in the future. Assuming that electricity may be sold for USD 0.07/kWh (ENTRO, 2007) and adopting a 3% annual discounting rate, the results (Table 5-4) indicate that during the filling stage, the 10% filling strategy produces a discounted present value on the order of USD 528 million. This is substantially lower than the annual benefits resulting from the 25%, 6-year, and 8-year strategies across the filling stage, which are all approximately equal (~USD 575 million).

Table 5-4: Discounted present value (as of December 2014) of annual benefit (million USD) from hydroelectricity generated during the filling stage (unique to each strategy) and full projection period (2018 – 2034), assuming electric price of \$0.07/kW-hrs, 3% annual discounting rate and full hydroelectric generation after the reservoir is filled (15 000 GW-hrs annually). Calculations are based on the mean values of 100 simulations.

Filling Strategy \ Benefit		10%	25%	HASF	4-yr	6-yr	8-yr
during filling stage	mean	527.7	573.8	541.0	553.9	575.2	574.4
	max	644.9	619.5	623.5	632.8	638.9	625.1
2018 - 2034	mean	535.5	643.9	632.4	668.9	644.4	620.3

5.6 Discussion

No agreeable filling strategy for the reservoir behind the GERD has been established between the riparian countries, even though serious upstream and downstream impacts are at stake. Two major factors – projected streamflow and various filling strategies – are explored here.

Although the low frequency signals explain only 17% of the variance in the streamflow time series, their inclusion is important for projecting across the filling period, particularly considering the uncertainty (based on the 100 simulations) in time to fill, downstream releases, and hydropower generation. The uncertainty in these low frequency signals, however, has not been explicitly included; rather, a static approach is undertaken. Additional investigation into the effects on outcomes of interest of stochastically representing the low frequency signals in projections may be worthwhile.

Regarding filling, the 10% strategy is probably impractical given the excessively long time required to fill the reservoir. The 4-year absolute strategy lies at the other end of the spectrum considered here, with a rapid filling time, implying sharp reductions downstream, including the likely occurrence of no flow months, and is unlikely to be selected if the decision is to be mutual. The HASF strategy typically favors downstream countries regarding streamflow, but results in significant uncertainty in the time to fill. The three remaining strategies (25%, 6-year, and 8-year) fall somewhere in the middle and may be considered as compromise solutions. The 25% and 6-year strategies are nearly identical in terms of time to fill and hydroelectricity generated; however, the 25% strategy has a larger uncertainty in outcomes while the 6-year strategy suffers from the possibility of months with no flow releases. Thus, a hybrid combination may be

warranted, combining the percentage-based and absolute filling strategies to reduce uncertainty and eliminate no-flow months, or alternatively selecting a minimum required flow to be released per month. The goal here is not to be exhaustive but rather to be illustrative in how different filling policies compare and whether they may favor upstream or downstream countries. Note that gridded data are used due to their longer record length, which is desirable for low frequency signal analysis. Gridded data can cause overestimation of the low flows and underestimation of the high flows, even with a carefully calibrated WatBal model (Chapter 4). For instance, the zero release events under the absolute filling strategies and the time required to fill the reservoir are likely to be overestimated. However, the cumulative error in terms of annual volume is approximately -750 million m³ (figure not shown), which is small compared to the substantial natural inflow (approximately 46,000 million m³). Additionally, for a relative comparison between filling strategies, this effect should be minimal.

As discussed earlier, no attempt is made here to include post-filling operational strategies. This will clearly be important for the management of infrastructure in all three countries. Sudan may be the least affected, given its relatively small amount of current storage capacity. Coordination of the GERD and the Aswan Dam in Egypt, however, is crucial for maximization of regional net benefits. In fact, cooperation among the three riparian countries at all stages – construction, filling, and long-term management – is critical for project success and future regional development.

Chapter 6 Summary and Recommendations

Irregular distribution of fresh water in both time and space can cause excess and scarcity, floods and droughts, food and energy crises, water conflicts, and many other physical and social challenges. To efficiently utilize and manage water resources requires a thorough understanding of climate, hydrology, social and economic aspects in the spatial-temporal distribution of water resources, potential optimal responsive strategies, and associated social welfare. In this dissertation, the ability of hydroclimatic prediction to skillfully guide decision making for agricultural planning and the application of hydroclimatic projection to inform reservoir filling in Ethiopia are explored. Methodologies developed in this work are also generalizable and can be applied to other places – a broader impact than simply in Ethiopia.

Contributions of this work may be most relevant for developing countries, where water resources utilization and management may be lacking due to minimal infrastructure, technological and scientific support for applying advanced methodologies, and practical guidance to encourage adaptation. These countries are also often poor, further motivating the need for contributions from the scientific community and relevant institutions. Concurrently application of this scientific information requires understanding and promoting realistic considerations, reliable outcomes, solid conclusions, and collaborations with decision-makers to improve livelihoods in the developing countries.

Several potential extensions and recommendations of this work are suggested. First, regarding prediction, exploration of methods to disaggregate predictions into monthly or even daily time-series for direct application into sectoral models is warranted. Understanding the lead time that is

required for these applications and how that is coupled with prediction skill is also important. Additionally, which prediction variables (e.g. seasonal total, dry spells within a season, the number of flood events) are most useful or relevant is not necessarily always evident – to the researcher or decision-maker – thus requiring further investigation and iterative communication.

Second, evaluating how farmers may or may not adopt climate predictions and suggested actions warrants further attention. These practices are clearly influenced by risk tolerance, degree of communication, loss or gain of confidence in the prediction, or other factors at the individual and community scale. Additionally, how best to communicate and interpret uncertainty in predictions, regional market aspects, multi-market interactions, supply and demand, and other externalities remains an open question, requiring interdisciplinary collaboration to adequately address it.

Third, addressing reservoir operations policies for both short-term (daily, monthly, seasonal operational rules) and long-term (decadal-scale strategic planning) perspectives concurrently – considering predictive information and multi-decadal variability – requires further investigation. This includes understanding and delineating the optimal or preferred space of acceptable total benefits (or other outcome) for various operational policies and multiple objectives, such as hydropower, agriculture, ecology, recreation, navigation, dam safety and ecological impacts.

Bibliography

- ABDELHADY, D., AGGESTAM, K., ANDERSSON, D.-E., BECKMAN, O., BERNDTSSON, R., PALMGREN, K. B., MADANI, K., OZKIRIMLI, U., PERSSON, K. M. & PILESJÖ, P. 2015. The Nile and the Grand Ethiopian Renaissance Dam: Is There a Meeting Point between Nationalism and Hydrosolidarity? *Journal of Contemporary Water Research & Education*, 155, 73-82.
- ADAMS, R. M., HOUSTON, L. L., MCCARL, B. A., TISCARENO, M. L., MATUS, J. G. & WEIHER, R. F. 2003. The benefits to Mexican agriculture of an El Nino-southern oscillation (ENSO) early warning system. *Agricultural and Forest Meteorology*, 115, 183-194.
- ALLEN, R. G., PEREIRA, L. S., RAES, D. & SMITH, M. 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *FAO, Rome*, 300, D05109.
- ANDERSON, D., STOCKDALE, T., BALMADEA, M., FERRANTI, L., VITART, F., MOLteni, F., DOBLAS-REYES, F., MOGENSEN, K. & VIDARD, A. 2007. Development of the ECMWF seasonal forecast System 3. *ECMWF Technical Memoranda*, 503.
- ARJOON, D., MOHAMED, Y., GOOR, Q. & TILMANT, A. 2014. Hydro-economic risk assessment in the eastern Nile River basin. *Water Resources and Economics*, 8, 16-31.
- BADR, H. S., ZAITCHIK, B. F. & DEZFULI, A. K. 2015. A tool for hierarchical climate regionalization. *Earth Science Informatics*, 1-10.
- BARRETT, C. B. 1993. THE DEVELOPMENT OF THE NILE HYDROMETEOROLOGICAL FORECAST SYSTEM1. Wiley Online Library.
- BARTLE, A. 2002. Hydropower potential and development activities. *Energy Policy*, 30, 1231-1239.
- BASTAWESY, M. 2014. Hydrological Scenarios of the Renaissance Dam in Ethiopia and Its Hydro-Environmental Impact on the Nile Downstream. *Journal of Hydrologic Engineering*, 20, 04014083.
- BEKELE, F. 1997. Ethiopian Use of ENSO Information in Its Seasonal Forecasts. *Internet Journal of African Studies*.
- BISETEGNE, D., L. OGALLO & ININDA, J. 1986. Rainfall characteristics in Ethiopia. *Reprint from 1st Technical Conference on Meteorological Research in Eastern and Southern Africa*. Nairobi, Kenya.
- BLACK, E., SLINGO, J. & SPERBER, K. R. 2003. An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean SST. *Monthly Weather Review*, 131, 74-94.
- BLOCK, P. & GODDARD, L. 2012. Statistical and Dynamical Climate Predictions to Guide Water Resources in Ethiopia. *Journal of Water Resources Planning and Management*, 138, 287-298.
- BLOCK, P. & RAJAGOPALAN, B. 2009. Statistical–Dynamical Approach for Streamflow Modeling at Malakal, Sudan, on the White Nile River. *Journal of Hydrologic Engineering*, 14, 185-196.
- BLOCK, P. & STRZEPEK, K. 2010. Economic Analysis of Large-Scale Upstream River Basin Development on the Blue Nile in Ethiopia Considering Transient Conditions, Climate Variability, and Climate Change. *Journal of Water Resources Planning and Management*, 136, 156-166.
- BLOCK, P. & STRZEPEK, K. 2012. Power Ahead: Meeting Ethiopia's Energy Needs Under a Changing Climate. *Review of Development Economics*, 16, 476-488.
- BLOCK, P. J., FILHO, F. A. S., SUN, L. & KWON, H. H. 2009. A Streamflow Forecasting Framework Using Multiple Climate and Hydrological Models. *Journal of the American Water Resources Association*, 45, 828-843.

- BLOCK, P. J. & RAJAGOPALAN, B. 2007. Interannual Variability and Ensemble Forecast of Upper Blue Nile Basin Kiremt Season Precipitation. *J. Hydrometeorol*, 8, 327-343.
- BLOCK, P. J., STRZEPEK, K., ROSEGRANT, M. W. & DIAO, X. 2008. Impacts of considering climate variability on investment decisions in Ethiopia. *Agricultural Economics*, 39, 171-181.
- BROAD, K. & AGRAWALA, S. 2000. The Ethiopia food crisis--uses and limits of climate forecasts. *Science*, 289, 1693-1694.
- CAMBERLIN, P. 1997. Rainfall Anomalies in the Source Region of the Nile and Their Connection with the Indian Summer Monsoon. *J. Climate*, 10, 1380-1392.
- CAMBERLIN, P. & PHILIPPON, N. 2002. The East African March–May Rainy Season: Associated Atmospheric Dynamics and Predictability over the 1968–97 Period. *Journal of Climate*, 15, 1002-1019.
- CHEN, J., BRISSETTE, F. P., CHAUMONT, D. & BRAUN, M. 2013. Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. *Water Resources Research*, 49, 4187-4205.
- CLEMONS, K. S. 2008. Hyrdoelectric Dams: Transboundary Environmental Effects and International Law. *Fla. St. UL Rev.*, 36, 487-536.
- COELHO, C. A. S., PEZZULLI, S., BALMASEDA, M., DOBLAS-REYES, F. J. & STEPHENSON, D. B. 2004. Forecast Calibration and Combination: A Simple Bayesian Approach for ENSO. *Journal of Climate*, 17, 1504-1516.
- CONWAY, D. 2000. The Climate and Hydrology of the Upper Blue Nile River. *Geographical Journal*, 166, 49-62.
- CONWAY, D. 2005. From headwater tributaries to international river: observing and adapting to climate variability and change in the Nile basin. *Global Environmental Change*, 15, 99-114.
- CONWAY, D. & SCHIPPER, E. L. F. 2011. Adaptation to climate change in Africa: Challenges and opportunities identified from Ethiopia. *Global Environmental Change*, 21, 227-237.
- CRAVEN, P. & WAHBA, G. 1979. Smoothing noisy data with spline functions. *Numerische Mathematik*, 31, 377-403.
- DEGEFU, W. 1987. Some aspects of meteorological drought in Ethiopia. In: GLANTZ, M. H. (ed.) *Drought and Hunger in Africa*. Cambridge University Press.
- DIAO, X. & PRATT, A. N. 2007. Growth options and poverty reduction in Ethiopia – An economy-wide model analysis. *Food Policy*, 32, 205-228.
- DIAO, X., PRATT, A. N., GHAUTAM, M., KEOUGH, J., CHAMBERLIN, J., YOU, L., PUETZ, D., RESNICK, D. & YU, B. 2005. Growth options and poverty reduction in Ethiopia. International Food Policy Research Institute (IFPRI).
- DINKU, T., HAILEMARIAM, K., MAIDMENT, R., TARNAVSKY, E. & CONNOR, S. 2014. Combined use of satellite estimates and rain gauge observations to generate high-quality historical rainfall time series over Ethiopia. *International Journal of Climatology*, 34, 2489-2504.
- DIRO, G. T., BLACK, E. & GRIMES, D. I. F. 2008. Seasonal forecasting of Ethiopian spring rains. *Meteorological Applications*, 15, 73-83.
- DIRO, G. T., GRIMES, D. I. F. & BLACK, E. 2011a. Teleconnections between Ethiopian summer rainfall and sea surface temperature: part I—observation and modelling. *Climate Dynamics*, 37, 103-119.
- DIRO, G. T., GRIMES, D. I. F. & BLACK, E. 2011b. Teleconnections between Ethiopian summer rainfall and sea surface temperature: part II. Seasonal forecasting. *Climate Dynamics*, 37, 121-131.
- DIRO, G. T., GRIMES, D. I. F., BLACK, E., O'NEILL, A. & PARDO-IGUZQUIZA, E. 2009. Evaluation of reanalysis rainfall estimates over Ethiopia. *International Journal of Climatology*, 29, 67-78.
- DIXON, B. L. & SEGERSON, K. 1999. Impacts of increased climate variability on the profitability of Midwest agriculture. *Journal of Agricultural and Applied Economics*, 31, 537-549.
- DOORENBOS, J. & KASSAM, A. 1979. Yield response to water. *Irrigation and drainage paper*, 33, 257.

- EEPCO 2013. Grand Ethiopian Renaissance Dam. Addis Ababa, Ethiopia: Ethiopian Electric Power Corporation.
- EKLUNDH, L. & PILESJÖ, P. 1990. Regionalization and spatial estimation of ethiopian mean annual rainfall. *International Journal of Climatology*, 10, 473-494.
- ELAGIB, N. A. & ELHAG, M. M. 2011. Major climate indicators of ongoing drought in Sudan. *Journal of Hydrology*, 409, 612-625.
- ELSAIED, S. M., HAMED, K., ASFAW, G., SELESHI, Y., AHMED, A. E., DEYAB, D. H., YON, B., ROE, J. D. M., FAILER, E. & BASSON, T. 2013. International Panel of Experts for Grand Ethiopian Renaissance Dam. Addis Ababa, Ethiopia: IPoE.
- ENTRO 2007. Coordinated Investment Planning - Generation. *Eastern Nile Power Trade Program Study*. AfDB.
- ERKYIHUN, S. T. 2015. *Multi-Decadal Stochastic Streamflow Projection an Application to Water Resources Decision Making in the Colorado River Basin*. PhD, University of Colorado.
- FAO-UNESCO 1988. Soil map of the world, revised legend. *World Resources Report*, 60, 138.
- GAMACHU, D. 1977. Aspects of climate and water budget in Ethiopia. *Aspects of climate and water budget in Ethiopia*.
- GEBRELUEL, G. 2014. Ethiopia's Grand Renaissance Dam: Ending Africa's Oldest Geopolitical Rivalry? *The Washington Quarterly*, 37, 25-37.
- GIANNINI, A., BIASUTTI, M., HELD, I. M. & SOBEL, A. H. 2008. A global perspective on African climate. *Climatic Change*, 90, 359-383.
- GIANNINI, A., SARAVANAN, R. & CHANG, P. 2003. Oceanic Forcing of Sahel Rainfall on Interannual to Interdecadal Time Scales. *Science*, 302, 1027-1030.
- GISSILA, T., BLACK, E., GRIMES, D. I. F. & SLINGO, J. M. 2004. Seasonal forecasting of the Ethiopian summer rains. *International Journal of Climatology*, 24, 1345-1358.
- GLOBAL RUNOFF DATA CENTER, F. I. O. H. G., INTERNATIONAL HYDROLOGICAL PROGRAMME, U. N. E. S. & CULTURAL, O. 2001. Monthly Discharge Data for World Rivers (except former Soviet Union). Boulder, CO: Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory.
- GODDARD, L. & GRAHAM, N. E. 1999. Importance of the Indian Ocean for simulating rainfall anomalies over eastern and southern Africa. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 104, 19099-19116.
- GONG, X. & RICHMAN, M. B. 1995. On the Application of Cluster Analysis to Growing Season Precipitation Data in North America East of the Rockies. *Journal of Climate*, 8, 897-931.
- GRIFFITHS, J. 1972. Ethiopian highlands. *World survey of climatology*, 10, 369-388.
- HAMMER, G. L., HOLZWORTH, D. P. & STONE, R. 1996. The value of skill in seasonal climate forecasting to wheat crop management in a region with high climatic variability. *Australian Journal of Agricultural Research*, 47, 717-737.
- HAMMER, G. L., NICHOLLS, N. & MITCHELL, C. 2000. *Applications of seasonal climate forecasting in agricultural and natural ecosystems*, Springer Science & Business Media.
- HANSEN, J. W. 2002. Realizing the potential benefits of climate prediction to agriculture: issues, approaches, challenges. *Agricultural Systems*, 74, 309-330.
- HANSEN, J. W., MASON, S. J., SUN, L. & TALL, A. 2011. Review of seasonal climate forecasting for agriculture in sub-Saharan Africa. *Experimental Agriculture*, 47, 205-240.
- HARGREAVES, G. H. & SAMANI, Z. A. 1982. Estimating potential evapotranspiration. *Journal of the Irrigation and Drainage Division*, 108, 225-230.
- HARRIS, I., JONES, P. D., OSBORN, T. J. & LISTER, D. H. 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology*, 34, 623-642.

- HARTIGAN, J. A. 1975. *Clustering algorithms*, John Wiley & Sons, Inc.
- HURST, H. E. 1952. *The Nile*, London, Constable & Co.
- INES, A. V. M. & HANSEN, J. W. 2006. Bias correction of daily GCM rainfall for crop simulation studies. *Agricultural and Forest Meteorology*, 138, 44-53.
- JAIN, A. K., MURTY, M. N. & FLYNN, P. J. 1999. Data clustering: a review. *ACM Comput. Surv.*, 31, 264-323.
- JOLLIFFE, I. 2002. *Principal component analysis*, Wiley Online Library.
- JONES, J. W., HANSEN, J. W., ROYCE, F. S. & MESSINA, C. D. 2000. Potential benefits of climate forecasting to agriculture. *Agriculture Ecosystems & Environment*, 82, 169-184.
- KAHSAY, T. N., KUIK, O., BROUWER, R. & VAN DER ZAAG, P. 2015. Estimation of the transboundary economic impacts of the Grand Ethiopia Renaissance Dam: A computable general equilibrium analysis. *Water Resources and Economics*, 10, 14-30.
- KALNAY, E., KANAMITSU, M., KISTLER, R., COLLINS, W., DEAVEN, D., GANDIN, L., IREDELL, M., SAHA, S., WHITE, G. & WOOLLEN, J. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society*, 77, 437-471.
- KASSAHUN, B. 1987. Weather systems over Ethiopia. *Proc. First Tech. Conf. on Meteorological Research in Eastern and Southern Africa*. Nairobi, Kenya: UCAR.
- KATZ, R. W., BROWN, B. G. & MURPHY, A. H. 1987. DECISION ANALYTIC ASSESSMENT OF THE ECONOMIC VALUE OF WEATHER FORECASTS - THE FALLOWING-PLANTING PROBLEM. *Journal of Forecasting*, 6, 77-89.
- KIBAROGLU, A. & ÜNVER, I. O. 2000. An institutional framework for facilitating cooperation in the Euphrates-Tigris river basin. *International Negotiation*, 5, 311-330.
- KING, A. & BLOCK, P. 2014. An assessment of reservoir filling policies for the Grand Ethiopian Renaissance Dam. *Journal of Water and Climate Change*, 5, 233.
- KIRTMAN, B. P., MIN, D., INFANTI, J. M., KINTER, J. L., PAOLINO, D. A., ZHANG, Q., VAN DEN DOOL, H., SAHA, S., MENDEZ, M. P., BECKER, E., PENG, P., TRIPP, P., HUANG, J., DEWITT, D. G., TIPPETT, M. K., BARNSTON, A. G., LI, S., ROSATI, A., SCHUBERT, S. D., RIENECKER, M., SUAREZ, M., LI, Z. E., MARSHAK, J., LIM, Y.-K., TRIBBIA, J., PEGION, K., MERRYFIELD, W. J., DENIS, B. & WOOD, E. F. 2014. The North American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction. *Bulletin of the American Meteorological Society*, 95, 585-601.
- KORECHA, D. & BARNSTON, A. G. 2007. Predictability of June–September Rainfall in Ethiopia. *Monthly Weather Review*, 135, 628-650.
- KORECHA, D. & SORTEBERG, A. 2013. Validation of operational seasonal rainfall forecast in Ethiopia. *Water Resources Research*, 49, 7681-7697.
- KUMAR, P. & FOUFOULA-GEORGIU, E. 1997. Wavelet analysis for geophysical applications. *Reviews of Geophysics*, 35, 385-412.
- KWON, H.-H., LALL, U. & KHALIL, A. F. 2007. Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature. *Water Resources Research*, 43, n/a-n/a.
- LALL, U. & SHARMA, A. 1996. A Nearest Neighbor Bootstrap For Resampling Hydrologic Time Series. *Water Resources Research*, 32, 679-693.
- LANDMAN, W. A. & MASON, S. J. 1999. Operational long-lead prediction of South African rainfall using canonical correlation analysis. *International Journal of Climatology*, 19, 1073-1090.
- LATIF, M., DOMMENGET, D., DIMA, M. & GRÖTZNER, A. 1999. The role of Indian Ocean sea surface temperature in forcing east African rainfall anomalies during December-January 1997/98. *Journal of Climate*, 12, 3497-3504.

- LETSON, D., PODESTA, G. P., MESSINA, C. D. & FERREYRA, R. A. 2005. The uncertain value of perfect ENSO phase forecasts: Stochastic agricultural prices and intra-phase climatic variations. *Climatic Change*, 69, 163-196.
- LIM, E.-P., HENDON, H. H., HUDSON, D., WANG, G. & ALVES, O. 2009. Dynamical Forecast of Inter-El Niño Variations of Tropical SST and Australian Spring Rainfall. *Monthly Weather Review*, 137, 3796-3810.
- MAHER, A. 2013. *Egyptian politicians caught in on-air Ethiopia dam gaffe* [Online]. BBC News. Available: <http://www.bbc.com/news/world-africa-22771563> [Accessed Jan 25 2016].
- MANNING, C. D., RAGHAVAN, P. & SCHÜTZE, H. 2008a. *Introduction to information retrieval*, Cambridge university press Cambridge.
- MANNING, C. D., RAGHAVAN, P. & SCHÜTZE, H. 2008b. *Introduction to information retrieval*, Cambridge university press Cambridge.
- MARSHALL, G. R., PARTON, K. A. & HAMMER, G. L. 1996. Risk attitude, planting conditions and the value of seasonal forecasts to a dryland wheat grower. *Australian Journal of Agricultural Economics*, 40, 211-233.
- MASON, S. 1998. Seasonal forecasting of South African rainfall using a non-linear discriminant analysis model. *International Journal of Climatology*, 18, 147-164.
- MESSINA, C. D., HANSEN, J. W. & HALL, A. J. 1999. Land allocation conditioned on El Nino Southern Oscillation phases in the Pampas of Argentina. *Agricultural Systems*, 60, 197-212.
- METZGER, S., LATIF, M. & FRAEDRICH, K. 2004. Combining ENSO Forecasts: A Feasibility Study. *Monthly Weather Review*, 132, 456-472.
- MEZA, F. J., HANSEN, J. W. & OSGOOD, D. 2008. Economic Value of Seasonal Climate Forecasts for Agriculture: Review of Ex-Ante Assessments and Recommendations for Future Research. *Journal of Applied Meteorology and Climatology*, 47, 1269-1286.
- MJELDE, J. W., SONKA, S. T., DIXON, B. L. & LAMB, P. J. 1988. Valuing forecast characteristics in a dynamic agricultural production system. *American Journal of Agricultural Economics*, 70, 674-684.
- MOFA 2012. Nile: Sudan's support for the Grand Ethiopian Renaissance Dam project. Addis Ababa, Ethiopia.: Ministry of Foreign Affairs, A Week in the Horn.
- MULAT, A. G. & MOGES, S. A. 2014. Assessment of the Impact of the Grand Ethiopian Renaissance Dam on the Performance of the High Aswan Dam. *Journal of Water Resource and Protection*, 06, 583-598.
- MUTAI, C. C., WARD, M. N. & COLMAN, A. W. 1998. Towards the prediction of the East Africa short rains based on sea-surface temperature-atmosphere coupling. *International Journal of Climatology*, 18, 975-997.
- NMSA 1996. Climate and agroclimatic resources of Ethiopia. *NMSA Meteorological Research Report Series*. Addis Ababa, Ethiopia: National Meteorological Services Agency of Ethiopia.
- OMONDI, P., OGALLO, L. A., ANYAH, R., MUTHAMA, J. M. & ININDA, J. 2013. Linkages between global sea surface temperatures and decadal rainfall variability over Eastern Africa region. *International Journal of Climatology*, 33, 2082-2104.
- ORAM, P. Sensitivity of agricultural production to climatic change, an update. International Symposium on Climate Variability and Food Security in Developing Countries, New Delhi (India), 5-9 Feb 1987, 1989. IRRI.
- PACHAURI, R. K. & REISINGER, A. 2007. Contribution of working groups I, II and III to the fourth assessment report of the intergovernmental panel on climate change. *IPCC, Geneva, Switzerland*, 104.

- PALMER, T., ALESSANDRI, A., ANDERSEN, U. & CANTELAUBE, P. 2004. Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bulletin of the American Meteorological Society*, 85, 853.
- PARTHASARATHY, B., KUMAR, K. R. & MUNOT, A. A. 1993. Homogeneous Indian Monsoon rainfall: Variability and prediction. *Proceedings of the Indian Academy of Sciences - Earth and Planetary Sciences*, 102, 121-155.
- PATT, A., SUAREZ, P. & GWATA, C. 2005. Effects of seasonal climate forecasts and participatory workshops among subsistence farmers in Zimbabwe. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 12623-12628.
- RAFTERY, A. E., MADIGAN, D. & HOETING, J. A. 1997. Bayesian Model Averaging for Linear Regression Models. *Journal of the American Statistical Association*, 92, 179-191.
- REUTERS. 2013. Ethiopia studies on Nile dam fall short, Egypt says. *Reuters*.
- ROECKNER, E., OBERHUBER, J. M., BACHER, A., CHRISTOPH, M. & KIRCHNER, I. 1996. ENSO variability and atmospheric response in a global coupled atmosphere-ocean GCM. *Climate Dynamics*, 12, 737-754.
- RONCOLI, C. 2006. Ethnographic and participatory approaches to research on farmers' responses to climate predictions. *Climate Research*, 33, 81-99.
- RONCOLI, C., JOST, C., KIRSHEN, P., SANON, M., INGRAM, K. T., WOODIN, M., SOMÉ, L., OUATTARA, F., SANFO, B. J. & SIA, C. 2009. From accessing to assessing forecasts: an end-to-end study of participatory climate forecast dissemination in Burkina Faso (West Africa). *Climatic Change*, 92, 433-460.
- SAHA, S., NADIGA, S., THIAW, C., WANG, J., WANG, W., ZHANG, Q., VAN DEN DOOL, H. M., PAN, H. L., MOORTHY, S., BEHRINGER, D., STOKES, D., PEÑA, M., LORD, S., WHITE, G., EBISUZAKI, W., PENG, P. & XIE, P. 2006. The NCEP Climate Forecast System. *Journal of Climate*, 19, 3483-3517.
- SALMAN, S. M. A. 2013. The Nile Basin Cooperative Framework Agreement: a peacefully unfolding African spring? *Water International*, 38, 17-29.
- SALMAN, S. M. A. 2016. *Sudan: Renaissance Dam - Limelight On the Fourth Ministerial Meeting & Khartoum Document* [Online]. Khartoum: SudaNow. Available: <http://allafrica.com/stories/201601080844.html> [Accessed Jan 25 2016].
- SCHEPEN, A., WANG, Q. J. & ROBERTSON, D. E. 2012. Combining the strengths of statistical and dynamical modeling approaches for forecasting Australian seasonal rainfall. *Journal of Geophysical Research: Atmospheres*, 117, n/a-n/a.
- SCHWABACH, A. 1998. United nations convention on the law of non-navigational uses of international watercourses, customary international law, and the interests of developing upper Riparians. *Tex. Int'l LJ*, 33, 257.
- SEGELE, Z. T. & LAMB, P. J. 2005. Characterization and variability of Kiremt rainy season over Ethiopia. *Meteorology and Atmospheric Physics*, 89, 153-180.
- SEGELE, Z. T., LAMB, P. J. & LESLIE, L. M. 2009. 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. *Journal of Climate*, 22, 3396-3421.
- SEGELE, Z. T., RICHMAN, M. B., LESLIE, L. M. & LAMB, P. J. 2015. Seasonal-to-Interannual Variability of Ethiopia/Horn of Africa Monsoon. Part II: Statistical Multi-Model Ensemble Rainfall Predictions. *Journal of Climate*, 150129124820009.
- SELESHI, Y. & ZANKE, U. 2004. Recent changes in rainfall and rainy days in Ethiopia. *International journal of climatology*, 24, 973-983.
- SHANKO, D. & CAMBERLIN, P. 1998. The effects of the Southwest Indian Ocean tropical cyclones on Ethiopian drought. *International Journal of Climatology*, 18, 1373-1388.

- SHUKLA, S., ROBERTS, J., HOELL, A., FUNK, C. C., ROBERTSON, F. & KIRTMAN, B. 2016. Assessing North American multimodel ensemble (NMME) seasonal forecast skill to assist in the early warning of anomalous hydrometeorological events over East Africa. *Climate Dynamics*, 1-17.
- SINGH, A., KULKARNI, M. A., MOHANTY, U. C., KAR, S. C., ROBERTSON, A. W. & MISHRA, G. 2012. Prediction of Indian summer monsoon rainfall (ISMR) using canonical correlation analysis of global circulation model products. *Meteorological Applications*, 19, 179-188.
- SOLOMON, S., QIN, D., MANNING, M., MARQUIS, M. & AVERYT, K. 2007. Technical Summary in Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007 24 National Research Council. *America's Climate Choices*.
- SOLOW, A. R., ADAMS, R. F., BRYANT, K. J., LEGLER, D. M., O'BRIEN, J. J., MCCARL, B. A., NAYDA, W. & WEIHER, R. 1998. The value of improved enso prediction to US agriculture. *Climatic Change*, 39, 47-60.
- STONE, R. C., HAMMER, G. L. & MARCUSSEN, T. 1996. Prediction of global rainfall probabilities using phases of the Southern Oscillation Index. *Nature*, 384, 252-255.
- SUDANVISION. 2013. Millennium dam: A report on the Ethiopia millennium dam, its impact on Sudan. *Sudan Vision*.
- SUGAR, C. A. & JAMES, G. M. 2003. Finding the Number of Clusters in a Dataset. *Journal of the American Statistical Association*, 98, 750-763.
- SUTCLIFFE, J. V. & PARKS, Y. P. 1987. Hydrological modelling of the Sudd and Jonglei Canal. *Hydrological Sciences Journal*, 32, 143-159.
- SWAIN, A. 2011. Challenges for water sharing in the Nile basin: changing geo-politics and changing climate. *Hydrological Sciences Journal*, 56, 687-702.
- TADESSE, T. 1994. The influence of the Arabian Sea storms/ depressions over the Ethiopian weather. *Proc. Int. Conf. on Monsoon Variability and Prediction*. Geneva, Switzerland: World Meteorological Organization.
- TEUTSCHBEIN, C. & SEIBERT, J. 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, 456-457, 12-29.
- THAN, K. 2011. Ethiopia moves forward with massive Nile dam project. *National Geographic Daily News*.
- THORNDIKE, R. L. 1953. Who belongs in the family? *Psychometrika*, 18, 267-276.
- THORNTON, P. 2006. Ex ante impact assessment and seasonal climate forecasts: status and issues. *Climate research*, 33, 55-65.
- TIBSHIRANI, R., WALTHER, G. & HASTIE, T. 2001. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63, 411-423.
- TORRENCE, C. & COMPO, G. P. 1998. A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, 79, 61-78.
- UNITEDNATIONS 2013. World population prospects: The 2012 revision, key findings and advance tables. In: DEPT. OF ECONOMIC AND SOCIAL AFFAIRS, P. D. (ed.). New York.
- USBR 1964. Land and water resources of the Blue Nile basin: Ethiopia. In: RECLAMATION, U. S. B. O. & U.S. DEPT. OF THE INTERIOR, G. (eds.). Washington, DC: Main Rep. and Appendices.
- VERHOEVEN, H. 2011. *Black Gold for Blue Gold?: Sudan's Oil, Ethiopia's Water and Regional Integration*, Chatham House.
- VISTE, E. & SORTEBERG, A. 2013a. The effect of moisture transport variability on Ethiopian summer precipitation. *International Journal of Climatology*, 33, 3106-3123.
- VISTE, E. & SORTEBERG, A. 2013b. Moisture transport into the Ethiopian highlands. *International Journal of Climatology*, 33, 249-263.

- WARD JR, J. H. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58, 236-244.
- WASHINGTONPOST. 2013. Ethiopian parliament ratifies new Nile sharing agreement opposed by Egypt. *Washington Post*.
- WATERBURY, J. 2002. *The Nile Basin: National determinants of collective action*, New Haven, CT, Yale University Press.
- WHEELER, K. G., BASHEER, M., MEKONNEN, Z. T., ELTOUM, S. O., MERSHA, A., ABDO, G. M., ZAGONA, E. A., HALL, J. W. & DADSON, S. J. 2016. Cooperative filling approaches for the Grand Ethiopian Renaissance Dam. *Water International*, 1-24.
- WILKS, D. S. 2011. *Statistical methods in the atmospheric sciences*, Academic press.
- WILKS, D. S. & MURPHY, A. H. 1986. A Decision-Analytic Study of the Joint Value of Seasonal Precipitation and Temperature Forecasts in a Choice-of-Crop Problem. *Atmosphere-Ocean*, 24, 353-368.
- WOLDE-GEORGIS, T. 1997. El Niño and drought early warning in Ethiopia. *Internet J. Afr. Studies*, 2.
- WORLDBANK 2006. Managing water resources to maximize sustainable growth: A Country water resources assistance strategy for Ethiopia. Washington, DC.
- WORLDBANK 2013. Ethiopia data. Washington, DC.
- WWAP 2012. Managing Water under Uncertainty and Risk. *The United Nations World Water Development Report*. Paris: UNESCO.
- WWF. 2007. Nile: Longest river on earth. *World Wildlife Fund*.
- YATES, D. N. 1996. WatBal: An Integrated Water Balance Model for Climate Impact Assessment of River Basin Runoff. *International Journal of Water Resources Development*, 12, 121-140.
- YATES, D. N. & STRZEPEK, K. M. 1998. Modeling The Nile Basin under Climate Change. *Journal of Hydrologic Engineering*.
- ZEITOUN, M. & ALLAN, J. A. 2008. Applying hegemony and power theory to transboundary water analysis. *Water Policy*, 10, 3.