Managing climate risks in agriculture and hydropower by coupling prediction and multi-purpose reservoir models in the Blue Nile Basin, Ethiopia

By
Sarah O. Alexander

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

Master of Science
(Civil and Environmental Engineering)

at the
UNIVERSITY OF WISCONSIN-MADISON
2018
ACKNOWLEDGEMENTS

This work would not be possible without the guidance, collaboration, and support of Dr. Paul Block. His commitment to research and advising has led to both personal and professional growth through this research. Committee members Dr. Daniel Wright and Dr. Anita Thompson as well as anonymous reviewers of a manuscript in preparation provided insight and critical feedback that has elevated the caliber of the work.

The material presented in this thesis is based upon work supported by the National Science Foundation under Grant No. 1545874.

Finally, I would like to acknowledge many organizations and individuals for their collaboration, insight, and feedback, including: National Meteorological Agency of Ethiopia, Ethiopia Ministry of Water Resources, Abay Basin Authority, Girmachew Addisu, Dr. Shu Wu, PIRE colleagues at the University of Connecticut, Water Systems and Society research colleagues, the Water Resources Engineering group, friends, and family. Their support and engagement have greatly contributed to the success of the work.
ABSTRACT

The Blue Nile Basin in Ethiopia is a critical water resource for the nation and continent, subject to spatial and temporal variations in climate that adversely impact local communities. Extreme conditions, e.g. drought and flood, threaten livelihoods causing direct hydrologic, economic, and social implications. To reduce vulnerability and enhance economic security in the region, predictive information is coupled with operational reservoir models to issue advance prediction of dominant rainy season conditions and inform sectoral decision-making.

Advance predictions of seasonal precipitation may provide information to aid water resource management decisions in various sectors. Yet, a disconnect between the spatial scale upon which skillful predictions are issued and the sectoral decision-making scale renders current predictive information inadequate in many cases. This work explores season-ahead precipitation prediction skill for a local region in the Blue Nile basin, Ethiopia, to better understand how model structure may serve to provide more skillful and valuable predictive information to the end-user. Statistical downscaling of global dynamic and regional empirical models as well as development of a high resolution, locally-tailored statistical model indicate that model structure and prediction skill are inextricably linked. Statistical approaches specific to the local region show higher prediction skill at the sectoral decision-making scale compared with dynamic approaches, offering the potential to aid local communities in many regions that are currently vulnerable to highly variable spatial and temporal precipitation patterns.

Predictive hydroclimate information may be further enhanced by integrating with sectoral models, to provide actionable information for communities by explicitly addressing climate risks. In this work, local-scale, statistical streamflow forecasts are coupled with a multi-purpose reservoir model to explore benefits and tradeoffs to hydropower and agriculture production for the Finchaa – Amerti system in the Blue Nile basin of Ethiopia. Strategies to meet agriculture demand and maximize hydropower result in increased hydropower production and agriculture water allocation using this framework. The resulting increase in resilience and decrease in vulnerability across sectors provides a concrete example of how hydroclimate prediction can be translated to actionable information to guide water resources management and bridge the existing disconnect between prediction and decision-making scales.
# TABLE OF CONTENTS

Chapter 1: Introduction and Context
  - Background: Blue Nile basin, Ethiopia  
  - Case Study Regions  
  - Local Scale Sectoral Models  

Chapter 2: Local-scale Precipitation and Streamflow Prediction
  - Data and Forecast Verification Metrics
    - Data  
    - Forecast Performance Metrics  
  - Downscaling Global Dynamic and Regional Empirical Models
    - Dynamic Models  
    - Empirical Cluster Model  
  - Local Statistical Prediction Model
    - Model Framework  
    - Model Evaluation and Comparison  

Chapter 3: Coupled Forecast and Reservoir Models
  - Application of Precipitation Predictions for Reservoir Management
    - Framework for Coupling Seasonal Precipitation and Reservoir Volume  
    - Using the Local Statistical Model to Predict Reservoir Storage  
  - Coupling Streamflow Predictions with Operational Reservoir Models
    - Model of the Finchaa – Amerti System  
    - Prediction-coupled Operational Reservoir Model  
    - Measures for Assessment of Model Performance  
    - Evaluation and Prediction-coupled Reservoir Model Performance  
    - Opportunity for Future Re-operation of the Finchaa – Amerti System  

Chapter 4: Discussion and Conclusions  

References
CHAPTER 1: INTRODUCTION AND CONTEXT

Background: Blue Nile basin, Ethiopia

The Blue Nile basin (BNB) in Ethiopia is a critical water resource for the region and continent, serving as the largest tributary to the Nile River by providing over 70% of downstream flows at the Aswan Dam (Conway 2000, 2005; Siam and Eltahir 2017). A majority of the basin experiences strongly seasonal precipitation, with most falling during the Kiremt rainy season, June – September (Yates and Strzepek 1998). The relative wealth of water resources in the basin affords the potential for economic gains through irrigated agriculture and hydropower. Current agricultural practices, however, are primarily rain-fed and represent the dominant sector in Ethiopia; thus, precipitation is integral to sustaining the economy and livelihoods. Yet precipitation across the region remains subject to high spatial and temporal interannual variability (Conway 2000, 2005; Segele and Lamb 2005; Zhang et al. 2016). Previous work indicates that these interannual variations result from movement of the inter-tropical convergence zone (ITCZ), regional climate interactions, topographic influences, and large-scale climate teleconnections, such as the El-Nino Southern Oscillation (ENSO) (Bekele 1997; Block and Rajagopalan 2007; Camberlin 1997; Conway 2000; Diro et al. 2009; Gissila et al. 2004; Korecha and Barnston 2007; Segele et al. 2009; Segele and Lamb 2005; Zhang et al. 2016). Teleconnections to the Indian Ocean and other high pressure systems are also cited as contributing factors to interannual variation (Black 2005; Folland et al. 1986; Segele and Lamb 2005). The strong spatial and temporal variability in hydroclimate often results in relatively proximal local areas experiencing varying amounts or timing of precipitation throughout the season. Thus, hydroclimate conditions averaged at the regional scale – for which predictions are most often available – may not always be representative of the local, decision-making scale.

Season-ahead predictions from dynamic and statistical models have been previously investigated for regions in East Africa. Dynamic model approaches have largely focused on precipitation prediction at the regional scale, using either regional climate models (RCMs) or statistical methods to downscale global climate models (GCMs) for East Africa, however some studies indicate little additional value with RCM downscaling (Cheneka et al. 2016; Diro et al. 2012; Nikulin et al. 2017; Ogutu et al. 2017; Shukla et al. 2014a). For precipitation across Ethiopia specifically, dynamic
model predictions generally exhibit modest to poor skill (Gleixner et al. 2017; Jan van Oldenborgh et al. 2005; Jury 2014; Shukla et al. 2016). Statistical model approaches for precipitation prediction across East African locations typically employ oceanic and atmospheric climate variables and teleconnections directly (Camberlin and Philippon 2002; Funk et al. 2014; Hastenrath et al. 2004; Mwale and Gan 2005; Ntale et al. 2003; Philippon et al. 2002; Yeshanew and Jury 2007). Some statistical approaches use analog approaches based on large-scale climate variables for precipitation prediction (e.g. Obled et al. 2002). Statistical modeling approaches applied specifically in Ethiopia include national-level (e.g. Korecha and Barnston 2007) and sub-national scale approaches to capture heterogeneity of moisture transport processes and precipitation, particularly in the Ethiopian highlands (Bisetegne et al. 1986; Diro et al. 2011a; Eklundh and Pilesjö 1990; Gissila et al. 2004; Zhang et al. 2016). The National Meteorological Agency (NMA) of Ethiopia has issued sub-national precipitation forecasts across all of Ethiopia, conditioned primarily on the state of ENSO, since 1987 (Korecha and Sorteberg 2013). Diro et al. (2011b) use Pacific, Indian, and Atlantic sea-surface temperature anomalies as predictors with multiple linear regression and linear discriminant analysis techniques. Still others focus at the BNB scale, using nonparametric, statistical methods to generate ensemble precipitation forecasts (Block and Rajagopalan 2007), or k-means statistical methods to optimally determine clustered regions and provide tailored precipitation predictions for each cluster (Zhang et al. 2016). Even though these empirical models demonstrate skill at the cluster level, regressing predictions to higher resolution (approximately 11 km) has been shown to cause a significant deterioration in skill (Zhang et al. 2017). Thus, even though numerous sub-national to East African regional scale seasonal precipitation predictions have been developed, all with modest skill, a connection to expected conditions at the local scale is practically nonexistent, yet could serve to support agricultural, energy, environmental, or other sectoral decisions.

**Case Study Regions**

Two local regions within the BNB, Ethiopia, are selected for local hydroclimate prediction model assessment based on vulnerability to climate variability and potential to improve local economic security (Figure 1). The regions exhibit a strongly seasonal distribution of precipitation, with precipitation primarily across the JJAS months, during which cereal and subsistence crops are planted (e.g. long and short cycle grains, coffee, pulses, oilseeds). Portions of each region also
benefit from an installed reservoir, allowing for irrigated agriculture during a second cropping season in October – May, and potentially increasing the annual agricultural yield and household income through cash crops (e.g. short-cycle grains, corn, sugar cane, coffee; Eriksson 2013). Thus, predictive information has the potential to enhance rain-fed and irrigated agricultural management decisions.

**Figure 1.** Map of Blue Nile (Abay) basin, Ethiopia with the Koga and Finchaa basins shown in red and Medium Blue Nile Region (MBNR) outlined.

The Koga basin is located in the Mecha Woreda, West Gojam Zone of the Amhara state, approximately 35 km southwest of Bahir Dar. The basin transitions from mountainous areas upstream to relatively flat areas downstream; elevations across the basin range from 3000 to 1800 meters above sea level (m.a.s.l.) (Gebrehiwot et al. 2010; Reynolds 2013). Annual precipitation averages 1419 mm, with approximately 75% coming in the JJAS season, however JJAS precipitation varies notably year to year (Figure 2a; coefficient of variation of 12%). The Koga Dam and Irrigation Site was constructed in 2011 to provide large-scale irrigation to farming communities downstream in an effort to increase economic security in the region through irrigated agriculture during the dry season. Located on the Koga River, a tributary to the Gilgel Abay River that flows to Lake Tana, the site includes a dam and 83.1 million cubic meter (Mm³) capacity reservoir to capture river inflow, precipitation, and runoff from a 220 km² catchment area. A system of canals carries water from the reservoir to provide irrigation to approximately 7000 ha of
agricultural land (Eriksson 2013; Reynolds 2013). In the Koga basin, over 70% of the river flow is a result of precipitation during the JJAS rainy season, thus the ability to capture water for a second crop in the dry season provides significant benefits to the surrounding communities (Reynolds 2013).

The Finchaa watershed is located in the southern portion of the Abay basin and covers approximately 1318 km². The watershed is predominately rolling plateaus with elevations ranging
from 3100 to 2200 m.a.s.l. Average annual precipitation is 1442 mm, with 72% falling during the JJAS rainy season; like Koga, Finchaa experiences notable inter-annual variability in rainy season precipitation (Figure 2b; coefficient of variation of 9%). The Finchaa Dam was constructed in 1973 to foster economic growth in the region through hydropower and irrigated agriculture (Tefera and Sterk 2008). Located on the Finchaa River, the dam is fed by the Finchaa and Amerti Rivers, with a 1050.5 Mm$^3$ reservoir. The hydropower plant has a 134 Megawatt (MW) installed capacity from 4 turbines (Wilson 2007). Downstream releases are utilized at a sugarcane plantation and factory, fisheries, and as a local water source, providing irrigation for over 20,000 ha of agricultural land (Belissa 2016). Hydropower and agriculture demands are higher in the non-peak flow months of October – May, with less water demanded during the JJAS rainy season, allowing for reservoir filling (Belissa 2016). A dam on the Amerti River fills a 64.4 Mm$^3$ reservoir, with diversions released to the Finchaa reservoir providing additional water storage. In 2012, a dam was installed on the Neshe River with a hydropower plant consisting of 97 MW total capacity (Wilson 2007). While Neshe is now considered part of the Finchaa-Amerti system, releases from the Neshe reservoir flow to the Finchaa River downstream of the reservoir. Thus, Neshe will not be considered as part of this study.

**Local Scale Sectoral Models**

Season-ahead predictions of hydroclimate variables, such as precipitation or streamflow, have the potential to aid water resource management in many areas, yet challenges remain, including skill at the decision-making scale, prediction lead time, and cascading uncertainty. Multiple forecast frameworks for season-ahead prediction, both statistical and dynamical, have been proposed in recent studies (Block and Goddard 2012; Coelho et al. 2004; Goddard et al. 2001; Goddard and Hoerling 2006; Kirtman et al. 2014; Korecha and Barnston 2007; Saha et al. 2006; Schepen et al. 2012; Shukla et al. 2014a; b). A common critique of these frameworks is their focus at a larger, regional scale compared with the local scale at which water resource management decisions are based (Gilles and Valdivia 2009; Pennesi 2007; Plotz et al. 2017). Research on communication and uptake of forecasts repeatedly indicates that the disconnect between the scale of regional predictions and local management decisions is a driving reason for lack of uptake, and further, predictions at the local, decision-making scale have been shown to enhance value to farmers (Gilles and Valdivia 2009; Plotz et al. 2017; Ziervogel et al. 2010). This motivates evaluation in
Chapter 2 of local-scale season-ahead prediction skill conditioned on regional dynamical and empirical models as well as a local empirical model to understand scale effects, and essentially address the question: are skillful hydroclimate forecasts at the local, sectoral decision-making scale possible for enhanced value to the end-user?

For a hydroclimate forecast to have utility at the decision-making scale, model structures must be capable of capturing small-scale variations in climate specific to the local region in addition to larger-scale teleconnections (Plotz, Chambers and Finn 2017; Gilles and Valdivia 2009; Ziervogel and Opere 2010). Dynamic models are state of the art, capturing complex climate dynamics and physical interactions at the global scale, valuable for enhancing climate information in data limited areas and for development of larger scale predictions. The accessibility of these forecasts as well as the number of inter-connected climate variables they produce, lends them useful for many applications (Kirtman et al. 2014; Shukla et al. 2016). Yet, biases, uncertainties, the inability to capture smaller scale interactions, errors in parameter estimates or initial conditions, and other discrepancies may result in unrealistic representations of local climate and poor prediction skill at local scales (Block and Rajagopalan 2007; Lupo and Kininmonth 2013). Alternatively, statistical prediction methods utilize climate teleconnections as predictors, and are often invoked for prediction at smaller scales. While statistical models typically do not predict the full suite of variables or well represent large spatial domains, as do dynamical models, the advantage of explicitly using historical observations to condition relationships, without fully capturing complex climate dynamics, may be in capturing smaller-scale inter-annual factors that modulate variability on a local scale (Korecha and Sorteberg 2013; Zhang 2016). Therefore, a tradeoff may exist between dynamic model structures, perhaps more suited to large-scale operational decision-making, and statistical model structures, which may capture smaller-scale variations, depending on the prediction scale desired. To evaluate, Chapter 2 compares local scale precipitation predictions from downscaled dynamic and regional models and a local statistical model within the Blue Nile basin.

Predictive information at the seasonal scale, linked with sectoral models offers the possibility of altering planning and management decisions to avert climate variability risks for end-users. Coupling of seasonal scale hydroclimate predictions with reservoir operations and management is
one example of how predictive information can be translated to actionable information with value to users. Water resource management teams are increasingly interested in using predictive information as a tool for mitigating risks from climate variability. For example, the U.S. Army Corps of Engineers and other organizations are part of a research assessment titled Forecast Informed Reservoir Operations (FIRO) to investigate whether climate forecast information can be used to manage and guide reservoir operations in California (Jasperse et al. 2017).

Previous studies have coupled seasonal forecasts with reservoir models to assess the ability of using predictive information to guide operational decisions, yet many of these studies rely on dynamic models to provide the forecast information (Block et al. 2009; Crochemore et al. 2016; Faber and Stedinger 2001). Dynamic models capture complex climate dynamics at the global scale and often are easily accessible, rendering them useful for application to reservoir operations. For example, Gong et al. (2010) develop a framework to use a set of a priori simulations that rely on the historical record to represent future conditions, maintaining the rule curve structure while adapting for smaller perturbations in the Delaware River Basin. Faber and Stedinger (2001) use ensemble streamflow predictions from the National Weather Service and sampling stochastic dynamic programming to optimize reservoir operations, while Crochemore et al. (2016) use Global Climate Model (GCM) forecasts conditioned historical data to make best use of statistical and dynamical methods.

Application of seasonal statistic predictions to reservoir management is limited, yet some researchers have found that statistical predictions may increase reservoir performance (Gelati et al. 2014; Georgakakos 1989; Golembesky et al. 2009; Hamlet et al. 2002). Maurer and Lettenmaier (2004) explore the use of long-lead predictions for reservoirs on the main stem of the Missouri River finding that economic benefits to hydropower are possible. Kim and Palmer (1997) use linear regression methods to develop a seasonal snowfall runoff forecast for input into a Bayesian Stochastic Dynamic Programming, demonstrating improved release patterns with the forecast. Sankarasubramanian et al. (2009) generate a framework for using probabilistic seasonal forecasts with water contracts for integrated management. More recently, LaMontagne and Stedinger (2018) illustrate how statistically derived synthetic streamflow forecasts can be input to a simple reservoir optimization to assess hydropower benefits. Given the recent success in using statistical methods
for prediction to inform reservoir operations and management, Chapter three seeks to understand how an existing local-scale precipitation and streamflow forecast can be translated through a reservoir model to answer the question: *can local-scale statistical seasonal forecasts be coupled with reservoir simulations to optimize operational strategies and provide actionable information to inform water resources management?*

Local scale precipitation predictions are used to issue an advance prediction of reservoir level at Koga (due to data limitations in providing streamflow prediction directly), in order to provide indication of how much water will be available in the dry season. Advance information of whether the reservoir is expected to fill or not may prove helpful in decisions made prior to the dry season including: amount of land to prepare, seed allocation, etc., with the opportunity for model output to provide some indication of the degree to which the reservoir may not fill. Coupling of Finchaa river streamflow predictions with a prediction-coupled reservoir model is used to understand the benefits to competing water users under different operational strategies and explore options for future re-operation.
CHAPTER 2: LOCAL SCALE PRECIPITATION AND STREAMFLOW PREDICTION

Data and Forecast Verification Metrics

Data

The Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) product, at 0.05 degree resolution, blends satellite information and gauge stations to provide a globally gridded precipitation dataset for 1981 to the present (Funk et al. 2015). CHIRPS has been utilized in many studies, and particularly in cases where observational records are poor (Funk et al. 2015; Shukla et al. 2016). Validation and comparison of CHIRPS against similar products, including Climate Forecast System (CFS) version 2, Climate Prediction Center Unified Gauge-based Analysis (CPCU), European Centre for Medium-Range Weather Forecasts (ECMWF), and TRMM Multi-satellite Precipitation Analysis (TMPA 3B42 RT7), indicates superior performance based on mean absolute error, mean bias, and correlation metrics (Funk et al. 2015). In the BNB, CHIRPS data is highly correlated with local precipitation gauge data (Figure 3) and is adopted as a representative observational precipitation record for the available 37 years (1981-2017; spatially averaged across the Koga (11.05 – 11.35 N, 37.05 – 37.35 E) and Finchaa (9.11 – 9.42 N, 37.00 – 37.30 E) regions).
Figure 3. CHIRPS and gauge station annual precipitation data for two stations near the Koga case study region (correlation: $a = 0.77$, $b = 0.43$).

The North American Multimodal Ensemble (NMME) is a set of hindcast and real-time ensemble forecasts (up to 9 months in advance) from a suite of Coupled Global Climate Models (CGCMs) (Kirtman et al. 2014). NMME precipitation predictions across East Africa demonstrate more skill than any single model and climatology (Shukla et al. 2016). Here, eight CGCMs providing 99 total ensemble members are downscaled to examine seasonal precipitation prediction for the Koga region in Ethiopia. Models used in analysis (and number of ensemble members) are as follows: CMC1-CanCM3 (10), CMC2-CanCM4 (10), CCSM4 (10), GFDL-CM2pl-aer04 (10), GFDL-CM2p5-FLOR-A06 (12), GFDL-CM2p5-FLOR-B01 (12), NASA-GMAO (11) and NCEP-CFSv2 (24). The NMME compares closely with other dynamic models, including ECMWF and GPCC (Funk et al. 2015). NMME model hindcasts are available from the International Research Institute for Climate and Society’s Data Library for 1982-2017 at a monthly, 1.0 degree resolution (Kirtman et al. 2014). NMME seasonal total JJAS precipitation predictions for each year are averaged spatially for each local region.
Statistical prediction models use global and local climate variables as predictors, including: sea surface temperature, sea level pressure, surface air temperature, and precipitable water content. All climate data were retrieved from the National Center for Environmental Prediction’s National Center for Atmospheric Research (NCEP-NCAR) reanalysis project data in 2.5 by 2.5 degree resolution for years 1948 to present (Kalnay et al. 1996).

**Forecast Performance Metrics**

Model hindcasts are evaluated with Pearson correlation, Hit Score, Extreme Miss Score, and Rank Probability Skill Score metrics to compare performance (Regonda et al. 2006). Here, categories are defined as below-normal (B), normal (N), and above-normal (A), with 33% of observations falling into each category (climatological distribution).

The Hit Score (Barnston 1992) is calculated as follows (equation 1):

\[
\text{Hit Score} = \frac{\sum \text{Hit}_A, \text{Hit}_N, \text{Hit}_B}{n} \times 100, \tag{1}
\]

where \(\sum \text{Hit}_A, \text{Hit}_N, \text{Hit}_B\) represents the sum of years where the median predicted ensemble value and observed value from the hindcast fall into the same category, and \(n\) represents the total number of years in the timeseries. Misses represent years where the predicted category is incorrect, and are particularly concerning when the prediction is off by more than one category (i.e. prediction of above-normal when below-normal was observed, \(\text{Miss}_{A(B)}\) or vice-versa), constituting an “extreme miss” (equation 2):

\[
\text{Extreme Miss Score} = \frac{\sum \text{Miss}_{A(B)} \text{Miss}_{B(A)}}{n} \times 100, \tag{2}
\]

where \(\sum \text{Miss}_{A(B)}\), \(\text{Miss}_{B(A)}\) represents the number of years deterministic prediction values are off by more than one category and \(n\) represents the total number of years in the timeseries.
Rank Probability Skill Score (RPSS) provides the categorical skill of an ensemble forecast with respect to a reference forecast (e.g. climatological distribution with A, N, B categories). The Rank Probability Score (RPS) is the average of the squared difference between the cumulative probability of the forecast and observations (equation 3):

\[ RPS = \sum_{i=1}^{n} (CP_{fct_i} - CP_{obs_i})^2, \]  

where \( CP_{fct_i} \) (\( CP_{obs_i} \)) represents the cumulative probability of the forecast (observations) through category \( i \), and \( n \) represents the total number of categories. The RPSS is then (equation 4):

\[ RPSS = \left[ 1 - \frac{RPS_{forecast}}{RPS_{climatology}} \right] \times 100, \]  

where \( RPS_{forecast} \) is the RPS for the forecast and \( RPS_{climatology} \) is the RPS for climatology. An RPSS of 100% indicates perfect forecast skill, with positive (negative) values indicating the forecast is more (less) skillful than climatology. The median RPSS value across all hindcast years is reported.

**Downscaling Global Dynamic and Regional Empirical Models**

**Dynamic Models**

Numerous approaches for dynamically and statistically downscaling dynamical model output are available, but not reviewed here. For this study, a quantile mapping approach to remove quantile dependent biases based on cumulative distribution functions of predicted and observed precipitation is adopted using drop-one-year cross-validation (Maraun 2013). Other approaches for downscaling and bias-correction exist, such as linear interpolation, spatial disaggregation, or downscaling with regional climate models, yet quantile mapping is selected here due to its simplicity and low computational requirements (Zhao et al. 2017). Downscaling of NMME precipitation predictions using model output statistics was also explored, yielding similar results. Statistically corrected ensemble NMME JJAS precipitation predictions issued June 1 for Koga (Figure 4, Table 1) illustrate relatively poor skill, indicating that the dynamic models are not well
capturing variability in seasonal precipitation at local scales. Additionally, predictions in some years miss by two categories (an “extreme miss”; Table 1). Although quantile mapping, or similar, is frequently applied bias-correction and post-processing (Bogner et al 2016), some argue that these methods are incapable of bridging the scale mismatch (Maraun 2013) as inconsistencies aside from model errors are introduced due to the inability to properly capture small-scale variability. Inadequate representation of smaller scale variabilities could account for the lack of skill in the bias-corrected NMME local predictions here and may warrant exploration of other bias-correction techniques.

![Figure 4](image)

**Figure 4.** Bias-corrected JJAS NMME ensemble averaged precipitation predictions (blue line), NMME ensembles (n = 99; boxes) and observations from CHIRPS (black line) for the Koga (a, correlation = 0.21) and Finchaa (b, correlation = 0.41) regions, 1982-2017.


**Table 1.** Deterministic and categorical verification metrics for bias-corrected NMME dynamic (1982-2017) and regional empirical (1982-2011) model JJAS precipitation predictions, issued June 1 for the Koga and Finchaa regions.

<table>
<thead>
<tr>
<th>Model</th>
<th>Local Region</th>
<th>Pearson Correlation</th>
<th>Rank Probability Skill Score (%)</th>
<th>HIT Score, % (# years)</th>
<th>Extreme Score, % (# years)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Downscaled Dynamic (NMME)</strong></td>
<td>Koga</td>
<td>0.21</td>
<td>-4.4</td>
<td>22.2 (8)</td>
<td>16.7 (6)</td>
</tr>
<tr>
<td></td>
<td>Finchaa</td>
<td>0.41</td>
<td>28</td>
<td>38.9 (14)</td>
<td>8.33 (3)</td>
</tr>
<tr>
<td><strong>Empirical Cluster Regression</strong></td>
<td>Koga</td>
<td>0.33</td>
<td>9.5</td>
<td>37.9 (11)</td>
<td>24.1 (7)</td>
</tr>
<tr>
<td></td>
<td>Finchaa</td>
<td>0.53</td>
<td>-11</td>
<td>41.4 (12)</td>
<td>6.9 (2)</td>
</tr>
</tbody>
</table>

**Empirical Cluster Model**

Zhang et al. (2017) propose an empirical seasonal prediction model for the BNB using a k-means clustering technique to assign homogeneous precipitation regions in an effort to decrease the larger regional prediction scale adopted by the NMA of Ethiopia (Korecha and Sorteberg 2013; Zhang et al. 2017). Once clusters are defined, the cluster level prediction is regressed to local regions within the cluster using a principle component regression model (PCR) framework (Zhang et al. 2017). Local predictions of JJAS seasonal precipitation issued on June 1 demonstrate increased correlation skill over NMME forecasts (Figure 5, Table 1) yet remain modest. Prediction correlation skill varies across clusters and is particularly poor along cluster boundaries (Korecha and Sorteberg 2013; Zhang et al. 2017). With the possible exception of Hit Score favoring the empirical cluster model approach, the remaining performance metrics do not show marked difference with NMME performance (Table 1). This moderate skill performance warrants investigation of alternative approaches, particularly those that are locally-tailored.
Figure 5. Bias-corrected JJAS empirical precipitation predictions (blue line) and observations from CHIRPS (black line) for the Koga (a, correlation = 0.33) and Finchaa (b, correlation = 0.53) regions, 1983-2011.

Local Statistical Prediction Model

Model Framework

A number of climate variables potentially modulating local inter-annual precipitation and streamflow through regional and global teleconnections are identified and evaluated as possible predictors of seasonal precipitation/streamflow by correlating gridded climate variables with observed JJAS precipitation (CHIRPS) or historical streamflow record at the case study location (e.g. Figure 6). Potential predictors include sea surface temperature (SST), sea level pressure (SLP), geopotential height (GH), meridional and zonal wind (MW), precipitable water content (PWC), and surface air temperature (SAT). Climate variable regions with a mean leave-one-year-out cross-validated correlation with observed JJAS precipitation at the 90% statistically significant level are retained for possible inclusion into the prediction model. Various lead times (forecast issue dates) are assessed (January 1, April 1, and June 1) for the precipitation predictions, using
only climate variable states prior to that date to better understand how prediction skill may change with longer lead time (Table 2). In addition to these climate variables, past studies indicate that temporal SST gradients, or tendencies, may also boost predictive skill (Diro et al. 2011a; Funk et al. 2014). To investigate, each statistically significant SST region is spatially averaged and all combinations of temporal SST tendencies from October – May are correlated with observed JJAS precipitation. Spatial SLP tendencies representing the North-Atlantic Oscillation and Indian Monsoon are also explored. Model predictors retained for each forecast issue date are presented in Table 2.

Figure 6. Correlation between Koga JJAS total precipitation from CHIRPS and global May sea-surface temperatures. Red boxes represent regions in the Indian Ocean (left) and Tropical Pacific Ocean (right) showing higher correlations.
Principle Component Regression (PCR), a common statistical method to reduce dimensionality and multicollinearity among large sets of data (Delorit et al. 2017; Lins 1985; Zhang et al. 2016, 2017), is adopted for the prediction model framework. PCR first transforms a set of variables (predictors) into uncorrelated, orthogonal principle components (PCs), ordered in terms of the amount of variance explained in the dataset. PCs explaining more than 10% of the overall dataset variance are retained, following Kaiser’s rule as in other studies (Delorit et al. 2017; Kaiser 1960; Zwick and Velicer 1986). In the second step, multiple-linear regression is performed using all retained PCs as predictors, as in equation 5:

$$\hat{y}_t = \beta_0 + \beta_1 PC_1 t + \beta_2 PC_2 t + \cdots + \beta_t PC_t \text{ for } t = 1 \text{ to } n,$$

$$e_t = \hat{y}_t - y_t,$$
where, $PC_t$ represents the selected PC at a given time step, $\beta$ values are the fitted regression model coefficients, $\hat{y}_t$ is the predicted value of JJAS precipitation/streamflow annually, and $e_t$ represents the residuals (predicted minus observed $y_t$ values) for each time step ($t$) from equation (5). To evaluate historical performance (had the prediction model been in place), a hindcast is performed in which, for each year $t = 1 \text{to} n$, the model is fitted in a leave-one-year-out cross-validation mode, providing a single (deterministic) JJAS precipitation/streamflow prediction for each observed year (Block and Rajagopalan 2007). Model residuals, $e_t$, from all hindcast years are fitted to a normal distribution with mean zero, using leave-one-year-out cross-validation, and added to the deterministic prediction values to form a distribution of predicted JJAS precipitation/streamflow for each year (Delorit et al. 2017; Helsel and Hirsch 1992; Zhang et al. 2016, 2017). The reliability of the probabilistic ensemble was verified using reliability diagrams (not shown; Hsu and Murphy 1986).

Using this PCR approach, JJAS precipitation predictions are developed for the forecast issue dates of interest (June 1st, April 1st, and January 1st; and JJAS streamflow predictions for June 1st). Model combinations for each lead time are compared using the Generalized Cross Validation (GCV) score, balancing model error and number of predictors by rewarding reduction in prediction error and penalizing overfitting (Block and Rajagopalan 2007; Craven and Wahba 1979). GCV is used in model selection by choosing a subset of statistically significant climate variable predictors, applying PCR, and then calculating the GCV using equation (7) below:

$$GCV = \frac{\sum_{t=1}^{N} e_t^2}{(1 - \frac{m}{N})},$$

where $e_t$ is the residual, $N$ is the number of observations, and $m$ is the number of predictor variables in the model. The model with the lowest GCV value for each lead time is retained.

Seasonal forecasts provide important indication of inflow volume during the dominant rainy season, yet monthly streamflow forecasts may also be informative for water resource management decisions. Following the PCR framework, the JJAS streamflow forecast is disaggregated to monthly streamflow predictions using an analog approach. Given the JJAS prediction, a year with
the closest total JJAS streamflow is chosen from the historical record and the monthly distribution of inflow is calculated. The predicted year is disaggregated to monthly predictions using the distribution from the analog year. Since month-to-month streamflow persistence is relatively strong following the JJAS season, monthly predictions for October, November, and December were obtained by persisting the predicted inflow value forward. Inflow during January – May is relatively consistent across the historical record, thus climatology is assumed for these monthly streamflow values. Methods result in JJAS seasonal prediction and a monthly streamflow prediction across all years.

The final set of predictors (Table 2) for the various leads for each region confirms many climate teleconnections that influence moisture transport to the region, further warranting their inclusion. For example, Pacific Ocean SSTs representing ENSO have been shown to influence JJAS seasonal precipitation in Ethiopia (Camberlin 1997; Diro et al. 2011b; a; Elagib and Elhag 2011; Segele and Lamb 2005). SSTs from the Indian and Atlantic Ocean basins also affect regional (Sahelian) precipitation on inter-decadal timescales (Folland et al. 1986; Giannini et al. 2008; Hoerling et al. 2006; Lu and Delworth 2005). Regional pressure systems such as the St. Helena, Mascarene, and Azores Highs have all been linked to JJAS moisture in Ethiopia (Black et al. 2003; Goddard and Graham 1999; Latif et al. 1999; Segele and Lamb 2005; Shanko and Camberlin 1998; Viste and Sorteberg 2013). Variations in Ethiopian climate have been strongly linked to the migration of the ITCZ and pressure changes resulting from the Indian Monsoon, which can be captured through SLP and local atmospheric variables such as PWC and SAT (Viste and Sorteberg 2013). Further, migration of the ITCZ results in an inverse pressure relationship between the summer and winter Indian Monsoon (Yancheva et al. 2007), thus a spatial wintertime SLP index across the Indian Ocean may serve as a precursor to the expected strength of the summertime monsoon during the JJAS season. In general, SST persistence is useful for inclusion in the model at most lead times, and addition of more transient SLP and atmospheric variables provides valuable predictive information at lead times closer to the season of interest.

**Model Evaluation and Comparison**

As expected, performance metrics typically decrease with increased lead time prior to the start of the season of interest (Table 3). For example, cross-validated predictions for Finchaa precipitation
(Figure 7) correlate with CHIRPS at 0.59 for a forecast issued on June 1 (start of season), while correlations decrease to 0.54 and 0.33, respectively, for April 1 and January 1 forecast issue dates. Although longer-lead predictions show a decrease in skill, predictive information provided as early as January may inform agricultural decisions that are made early in the calendar year (e.g. 22 years, 59% of the record, January and June predict the same category and 11 of those years categorical prediction aligns with observed). Finchaa streamflow predictions correlate with observed streamflow at 0.50 for a forecast issued June 1st (Figure 8). While a forecast issued June 1 incorporates SST, SLP, and atmospheric climate variables as predictors, more transient atmospheric and SLP variables not surprisingly illustrate less skill (and are not included) for the April and January forecasts given the longer lead times prior to the season of interest. Notably, RPSS is positive for all lead times, indicating an improvement over climatology.

For Koga, a June 1 prediction uses the moderately developed, although not yet fully known, ENSO state and includes SST, SLP, and atmospheric variables as predictors, resulting in a cross-validated correlation of 0.57. Correlations drop substantially, however, for the April 1 forecast (0.13), likely attributable to the well-documented “spring barrier” during which ENSO resets for the following seasons (Chen et al. 2017; Clarke and Van Gorder 1999; Jin et al. 2008; Webster and Yang 1992). Model skill at a January 1 forecast lead relies on SST climate signals from the northern hemisphere fall and winter, which prove skillful for years when these signals persist into the following summer; cross-validated predictions (Figure 7) correlate with CHIRPS at 0.52. RPSS values for all lead times are positive. In 20 years, 54% of the record, January and June predict the same category (aligning with the observed category 14 years), illustrating the possible value of longer-lead predictions to inform crop decisions made early in the year.
Figure 7. Local statistical model JJAS precipitation predictions (median = blue line, ensemble distribution = boxes) and JJAS precipitation from CHIRPS (black line) for the Koga (a – June forecast date, b – April, c – January) and Finchaa (d – June, e – April, and f – January) regions.

Figure 8. Local statistical model JJAS streamflow predictions (median = blue line, ensemble distribution = boxes) and observed JJAS streamflow (black line) for the Finchaa (a – June forecast date) and Amerti (b – June forecast data) Rivers.
For comparison, correlations of NMME predictions and CHIRPS precipitation for Koga (0.21, 0.15, and 0.02 for forecast issue dates of June 1, April 1, and January 1, respectively) and Finchaa (0.41, 0.04, and 0.02 for June 1, April 1, and January 1, respectively) are notably less for all lead times. These cumulative results suggest the superior performance of the local statistical model in predicting local scale variability in precipitation.

Comparison of NMME and local statistical predictions at additional spatial and temporal scales further illustrates the intricate link between model structure and prediction skill. Temporally, both downscaled NMME and the local statistical model are less skillful in predicting a single month of
the JJAS season than in predicting the JJAS seasonal total precipitation, based on deterministic Pearson correlation and categorical (using the full prediction ensemble) RPSS metrics (Figure 9a). This increasing skill with temporal aggregation is not unexpected and has been posited by others (e.g. Bohn et al. 2010). Perhaps more interestingly, performance metrics indicate that the local statistical model is superior across all temporal prediction scales (Figure 9a).

Figure 9. Comparison of correlation (left y-axis, black) and RPSS (right y-axis, blue) prediction skill for Koga from the NMME (o) and the local statistical model (*) for temporal (a; 1, 2, and 4-month) and spatial (b; small – Koga basin, 1,110 km², medium – MBNR, 49,284 km², and large – BNB, 308,025 km²) prediction periods. All forecasts issued June 1.

Considering JJAS precipitation predictions at different spatial scales, the local statistical model is superior at the small scale, while at the large scale NMME appears more suitable (Figure 9b). However, NMME skill may increase at the medium and small scale with alternative downscaling
methods and statistical model prediction skill may increase if model structures are tailored to the respective regions.

Based on this comparison, at a regional or basin scale, NMME prediction models may suffice, however, at the local level, a tailored statistical prediction framework may be warranted. Skillful prediction, however, does not necessarily directly translate into improved sectoral prediction and decision-making.
CHAPTER 3: COUPLING PREDICTIONS WITH RESERVOIR MODELS

Application of Precipitation Predictions to Reservoir Management

Framework for coupling seasonal precipitation and reservoir volume

Coupling skillful precipitation predictions at the local scale with sectoral operations and management may enhance local economic and environmental benefits. For example, a reservoir at Koga allows for irrigated agriculture during a second cropping season (October – May), yet, given interannual variability in precipitation, reservoir volumes at the end of the rainy season vary year to year, which affects agricultural land preparation, seed procurement, planned water allocation, etc. These decisions are made in advance of the end of the rainy season (October), thus the June 1 precipitation predictions and reservoir state are used to predict October reservoir volume.

General reservoir characteristics, monthly water release data, and reservoir stage data were compiled from multiple sources and provided by project partners (Birhanu et al. 2014; Mott MacDonald 2006). Following previous research in East Africa, monthly potential evapotranspiration was estimated using radiation and temperature through the Hargreaves method (Block and Goddard 2012; Droogers and Allen 2002; George H. Hargreaves and Zohrab A. Samani 1985; Hargreaves and Samani 1982). The following reservoir water balance (equation 8) was developed for simulation:

\[
V(t) = V(t-1) + P(t) \times SA(t) - ET(t) \times SA(t) + I(t) - R(t),
\]  

where for each monthly time step \( t \), \( V \) is reservoir volume (\( m^3 \)), \( P \) is total precipitation (\( m \)), \( ET \) is evapotranspiration (\( m \)), \( SA \) is reservoir surface area (\( m^2 \)), \( I \) is inflow (\( m^3 \)), and \( R \) represents water released from the reservoir (\( m^3 \)). Minimal information of initial hydrological conditions is available for the Koga basin, most notably soil moisture conditions and groundwater interactions, and are therefore not directly modeled as individual terms, similar to previous studies (Birhanu et al. 2014; Mott MacDonald 2006; Reynolds 2013). However, inflow values effectively represent direct flow from the Koga River as well as other fluxes (groundwater, soil moisture, infiltration, etc.) not explicitly accounted for in this equation due to data limitations. Initial hydrologic conditions may be less critical for long-lead predictions (greater than one month) and small-sized basins, thus implicit inclusion of these terms with ‘inflow’ is assumed for this application (Li et
al. 2009; Shukla and Lettenmaier 2011). Observations of reservoir volume, precipitation, and releases, along with estimated evapotranspiration, were used to determine inflow to the reservoir during the JJAS season for each year of operation (2013-2017). Based on this, a direct relationship between the total JJAS season precipitation-evapotranspiration difference and volume of inflow to the reservoir during JJAS was also developed (equation 9):

\[ I_{JJAS} = C_1 \times (P_{JJAS} - ET_{JJAS}) - C_2, \]  

where \( I_{JJAS} \) is the volume of inflow during JJAS and \( P_{JJAS} - ET_{JJAS} \) represents the total JJAS precipitation less the total evapotranspiration volume during the JJAS season. Coefficients \( C_1 \) and \( C_2 \) have mean values of 0.14 (0.08 – 0.21) and 60 (0.83 – 120) based on a drop-one-year cross-validation fit of the available observed record (mean coefficients are used for hindcast). Although this does not explicitly account for base flow, inflow from the Koga River is dominated by rainfall-runoff processes during this season. October 1 reservoir volumes are then the result of observed May reservoir volumes plus \( I_{JJAS} \) values. Although verification is limited to the years 2013-2017, simulated October reservoir volumes do match observed reservoir volumes in terms of full versus not full (reservoir filled in all years except 2015; Figure 10). Using CHIRPS data for 1981-2017 this relationship is extended to simulate the expected volume of JJAS inflow for historical years and October 1 reservoir volumes, contingent on average May storage volumes from 2013-2016 (Figure 10; black line). This hindcast simulation suggests that had the reservoir been in place for the period 1981-2017, it would have filled in approximately 62% of the years (greater than 50% probability). Thus, an October reservoir state less than full is not unusual.
Figure 10. Hindcast (1981-2017) of simulated October Koga reservoir levels (black), mean climatology (blue) and distribution of predicted October 1 reservoir volume (boxes) using local statistical model precipitation predictions issued June 1. Full reservoir storage level (83.1 million cubic meters) shown in red. Predicted probability of filling (%) shown above each year (green = hit, red = miss).

**Using the local statistical model to predict reservoir storage**

Probabilistic local statistical model precipitation predictions issued on June 1 for the period 1981-2017 are input into the framework to provide a hindcast of predicted October reservoir volumes. JJAS inflow volumes (Equation 9), based on precipitation predictions and average evapotranspiration values, were then added to an average May reservoir volume to illustrate how the framework can provide a distribution of predicted October reservoir volumes (Figure 10). Leave-one-year-out cross-validation is used for predictor selection, PCR model calibration, and error distributions for the precipitation predictions; the inflow model is also constructed in leave-one-year-out cross-validation mode to provide reservoir volume predictions. While not a robust forecast informed reservoir simulation, the framework allows precipitation predictions to be translated into information regarding reservoir stage. For irrigated agriculture communities that rely on both precipitation and reservoir allocation, this simple translation coupled with the precipitation prediction may provide a more complete indication of the likely amount of water to be received during the coming season.
The predicted median October reservoir volume agrees categorically (fill or no fill) with the reconstructed October reservoir volumes in 27 years (73% of the timeseries). The JJAS reservoir state is correctly predicted (probability greater than 50%) on June 1 for four years in which observational data is available (2013-2017) based on the median forecast (very close one year). Percentage probability of filling for each year shows how ensembles can be used to determine a prediction, where greater than 50% probability corresponds with a full reservoir (Figure 10). While a strictly categorical fill or no fill prediction was used for this example, ensemble distributions provide the opportunity to predict the approximate volume in reservoir for no fill scenarios. When climatological precipitation is used as the forecast (naïve benchmark; average May storage used for hindcast years, observed for operational years, to obtain climatological prediction), the reservoir state is always predicted to be full (Figure 10). While this happens to be correct in four of the five years, missing the 2015 drought may have serious consequences, particularly since the reservoir is expected to not fill frequently (approximately 1/6 of all years; Figure 10).

In years that the reservoir would not have filled, according to simulated ‘observations,’ the mean reservoir volume was 76.8 Mm$^3$ (6.3 Mm$^3$ short of full on average). For these same years, the average predicted reservoir volume was 84.2 Mm$^3$ (average deficit of 4.48 Mm$^3$) based on the median prediction. Further, the 2015 season was anomalously dry in the region and anecdotal evidence states the reservoir was just over half full by the end of the JJAS season. Results indicate October 2015 reservoir volumes of 57.0 Mm$^3$ and 63.4 Mm$^3$ for observed, median predictions. The predicted reservoir volumes and deficits during years that the reservoir does not fill compare remarkably well with the average volumes and deficits of the simulated observations. Full validation and assessment of the framework performance are limited given the lack of long observational records of inflow, operational decisions, and other data availability challenges. Nonetheless, an indication of forecast value and performance is still strongly inferable (Zhao et al. 2017; Turner et al. 2017).

**Coupling Streamflow Predictions with Operational Reservoir Models**

**Model of the Finchaa-Amerti Reservoir System**

A model of the Finchaa-Amerti reservoir system was created to simulate movement of water through the system, calculating hydropower generation and water allocation downstream. The
reservoir simulation uses historical inflow data, knowledge of current monthly release patterns, precipitation, evapotranspiration, as well as known characteristics of the Finchaa-Amerti system in an iterative water balance approach (equation 8). Minimal information regarding hydrologic conditions (e.g. soil moisture and groundwater interactions) is available for the region, and thus are omitted from the water balance calculation as in previous local studies in the BNB (Birhanu 2014; MacDonald 2006; Reynolds 2013).

Area, elevation, and volume relationships for both Finchaa and Amerti were generated based on available data for use in the reservoir simulation (Eastern Nile Power Trade Program Study). The Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) gridded precipitation product and estimation of evapotranspiration using the Hargreaves method are adopted, as in Chapter two (Funk et al. 2015; Hargreaves and Samani 1982). Iterations of the reservoir model were performed for calibration and to ensure that important constraints on the system (i.e. reservoir volume, elevation and area, power generation) were met. The reservoir model was validated using historical inflow records and current operations, comparison with results from a past study, anecdotal information, and known power generation for some historical years (Belissa 2016; Addisu 2018). The model of the reservoir system was used in the prediction-coupled operational reservoir model, below.

**Prediction-coupled Operational Reservoir Model**

The existing disaggregated statistical JJAS streamflow prediction and reservoir simulation model are coupled to investigate the skill of a prediction-coupled operational reservoir model to aid water resource decision-making under variable climate conditions. Four operational strategies are explored: A) current operations, B) meeting agricultural demands, C) maximizing hydropower production using local-scale streamflow predictions, and D) maximizing hydropower production using local-scale streamflow predictions and the Amerti reservoir as additional storage (i.e. multiple-reservoir). To couple the prediction and reservoir simulation models, an intermediate optimization function is developed with the goal of determining optimal releases for the prediction-coupled reservoir model at each timestep using single-objective linear optimization.
Forecasted monthly streamflow, precipitation and evapotranspiration climatology were input to the reservoir system simulation framework. At each monthly timestep, the forecasted Finchaa reservoir inflows for the following 12 months were fed to the optimization framework to determine optimal release strategies. The optimal release value for the current month was then input to the reservoir simulation to determine reservoir state (volume, surface area, elevation). Stepping forward, reservoir state from the previous month is updated based on observed inflow values before the next month is forecast and the process repeated. For example, in the beginning of May, forecasted inflow for the next year is optimized to determine the May release. In June, May reservoir state is updated with observed inflow information before using the 12-month forecast to optimize and determine the release value for June (Figure 11). At all timesteps, environmental river flows and water supply demanded downstream are required to be met. The model also requires storage at the end time period to equal initial storage for each optimization step, and input initial storage is assumed to be the cited normal operating level (536.4 Mm$^3$; Belissa 2016). Amerti is modeled primarily as a storage reservoir, only releasing water that would otherwise spill.

**Prediction-coupled Reservoir Model Framework**

![Prediction-coupled Reservoir Model Framework](image)

**Figure 11. Model framework for the prediction-coupled reservoir model.**

Multiple scenarios are input to the reservoir model to understand performance under varying operational strategies. First, the current static operational release pattern is simulated as a baseline
(Scenario A). Next, a static operational policy to meet agricultural demand is explored (Scenario B). Finally, the local-scale disaggregated statistical streamflow forecast for Finchaa (developed in Chapter two) is used in the prediction-coupled reservoir framework to investigate a dynamic release schedule with the objective of optimizing release for maximum hydropower generation using single-reservoir (Scenario C) and multi-reservoir (Scenario D) systems.

**Measures for Assessment of Model Performance**

Model performance is evaluated by identifying the degree to which simulated reservoir volumes and release schedules meet the needs of reservoir users, with benefit defined in terms of energy generation (hydropower) and volume of water allocated (agriculture). The range in monthly and yearly hydropower generation is quantified, as well as the total energy production over the 37-year timeseries. Further, performance during the most wet (1991) and dry (1983) years in the historical record provides indication of model performance during extreme conditions. Criteria of reliability, resilience and vulnerability have been used in multiple reservoir studies and shown to provide helpful information on performance of the system being modeled (Jain and Bhunya 2008). Taken together, these criteria are used to further assess model performance under varying scenarios.

Reliability is the probability of whether or not the system is in a satisfactory state, ranging from negative one (unsatisfactory) to positive one (satisfactory), and shown by the annual reliability metric (equation 11):

$$\text{Reliability} = 1 - \left( \frac{F_{\text{year}}}{T_{\text{year}}} \right)$$  \hspace{1cm} (11)

where $F_{\text{year}}$ refers to the number of years of failure and $T_{\text{year}}$ refers to the total number of years.

The resilience criterion describes how quickly a system is likely to rebound from being in a failed state. Ideally, a system would quickly respond from a state of failure to reach a satisfactory state in a timely manner. The resilience of the reservoir system is calculated as the inverse of the mean time that is spent in the unsatisfactory state, following Jain and Bhunya (2008), using equation 12:
Resilience = \left[ \frac{1}{N} \sum_{i=1}^{N} d(i) \right]^{-1}, \quad (12)

where \( N \) refers to the number of failure events, and \( d \) represents the duration of failure event, \( t \).

Vulnerability refers to the likely magnitude of a failure, should one occur, and is a measure of the severity of damage during a failed event. For this work, vulnerability was assessed in terms of the difference in water allocation (or energy production) demanded and provided, as in equations 13 and 14:

\[ Vuln_{ag} = \text{volume demand} - \text{volume allocated}, \quad (13) \]

\[ Vuln_{hydro} = \text{energy demand} - \text{energy generated}, \quad (14) \]

where \( Vuln_{ag} \) = vulnerability with respect to agriculture and \( Vuln_{hydro} \) = vulnerability with respect to hydropower. The water allocation demand for irrigated agriculture was defined as in previous work (Belissa 2016) and the desired monthly energy was taken as the current energy production during normal operations.

**Evaluation of Prediction-coupled Reservoir Model Performance**

The coupled prediction-reservoir model illustrates benefits to competing users under different operational Scenarios (A, B, C, D as introduced above). Modeling current operational strategies (Scenario A) based on a static rule curve developed using decades-long records of hydropower demand in the system results in moderate energy generation (22.0 – 65.4 GWh monthly, 21731 GWh total) and relatively low allocation for downstream users (6.55 – 40.0 Mm³ monthly and 9921 Mm³ total; Table 4). Performance of the Finchaa reservoir under current operational strategies yields low reliability (0.38 average) and resilience (0.15 average) and high vulnerability (6.68 average) criteria values, indicating that performance of the system with respect to both hydropower and irrigated agriculture is moderate at best (Figure 12). As shown in Figure 13a, there are many months in which benefits to hydropower are low and irrigation deficits high. While the current operational policies may have met past management plans, Ethiopia is rapidly
expanding hydropower production through re-operation of existing facilities and newly built infrastructure. Thus, a re-evaluation of production across the grid to determine improved operational strategies moving forward is warranted.

Table 4. Evaluation metrics for reservoir simulations for the prediction-coupled reservoir model, all scenarios and years 1981-2017.

<table>
<thead>
<tr>
<th>Model Scenario</th>
<th>Reservoirs Modeled</th>
<th>Forecast</th>
<th>Reliability</th>
<th>Resilience</th>
<th>Vulnerability</th>
<th>Energy (GWh)</th>
<th>Water Allocation (Mm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Current Operations</td>
<td>Finchaa</td>
<td>--</td>
<td>0.17</td>
<td>0.59</td>
<td>0.38</td>
<td>0.20</td>
<td>0.15</td>
</tr>
<tr>
<td>B. Meet Agriculture Demand</td>
<td>Finchaa</td>
<td>--</td>
<td>0.85</td>
<td>0.69</td>
<td>0.77</td>
<td>0.28</td>
<td>0.27</td>
</tr>
<tr>
<td>C. Hydropower Maximizing</td>
<td>Finchaa and Amerti</td>
<td>STAT</td>
<td>0.55</td>
<td>0.57</td>
<td>0.56</td>
<td>0.18</td>
<td>0.19</td>
</tr>
<tr>
<td>D. Hydropower Maximizing</td>
<td></td>
<td>STAT</td>
<td>0.61</td>
<td>0.67</td>
<td>0.64</td>
<td>0.22</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Note: Forecast - forecast, AG - agriculture, HYD - hydropower, AVG - average of both users, STAT - statistical, seasonal forecast. Number corresponds to scenario as introduced in the text.

Figure 12. Parallel line plot showing resilience, reliability, and vulnerability metrics for the operational strategies investigated using the prediction-coupled reservoir model framework. Note: vulnerability metric results have been standardized, and metrics reflect the average criteria value with respect to hydropower and agriculture.
Figure 13. Average monthly benefits to hydropower (dots) and water allocation deficits for irrigated agriculture (bars) with lines denoting 25th and 75th percentile (lines) for a) current operations, b) meeting agriculture demand, c) maximizing hydropower and d) maximizing hydropower with a multiple-reservoir scenario.

Comparing Finchaa reservoir monthly releases based on current operations and agriculture demands illustrates unmet allocation (Figure 14), on the order of 153 Mm$^3$ annually. Thus, Scenario B considers the performance of the Finchaa system re-operated to meet agriculture demands as often as possible (where demands defined as in Belissa 2016). Interestingly, results indicate greater benefit to both hydropower and agriculture under this scenario, as more water is passed through the turbines before allocation for irrigated agriculture. Additional water passed through the turbines, particularly in critical high flow months (September – November) results in increased net benefits to hydropower in addition to irrigated agriculture in Scenario B (Figure 13b). Average metrics across sectoral users display increased reliability (0.77; 49% improvement over current operations, Scenario A), resilience (0.27; 56% improvement) and decreased vulnerability (4.29; 64% improvement) of the system and show enhanced performance (Table 4). Total energy production (24745 GWh) and water allocation for agriculture (12148 Mm$^3$) are nearly 12.2% and
18.3% higher, respectively, than under current operational strategies, highlighting the benefit to each.

Figure 14. Comparison of Finchaa releases based on current operations (black) and agricultural demands (red; Belissa 2016); unmet demand shaded in grey.

Both Scenarios A and B utilize a static operational curve and do not consider expected upcoming conditions. Given substantial inter-annual climate variability within the BNB (coefficient of variation of rainy season precipitation is 9% near the Finchaa Reservoir), predicted inflow conditions – leading to dynamic operations – may prove valuable, particularly in dry years and those following.

To explore coupled prediction-reservoir operations (Scenario C), the disaggregated local-scale statistical streamflow prediction introduced in Chapter two is integrated with the Finchaa reservoir model with the objective of maximizing hydropower generation. A dynamic, or flexible, operational strategy (i.e. flexible reservoir rule curve) and the integration of forecast information serve to increase the performance of the reservoir system, as shown by moderately high reliability and resilience metrics and low vulnerability (Figure 12, Table 4). Important to note, comparison of scenario metrics in Figure 12 reflect the average with respect to hydropower and agriculture users, thus limited information on hydropower demand reduces relative skill of the prediction
informed scenario and improved demand information may alter results. Scenario C yields higher energy generation (range = 365 – 990 GWh annually; Table 4) and water allocation (range = 129 – 572 Mm³ annually; Table 4) than current operations (Scenario A; energy range = 237 – 624 GWh and water allocation range = 40.0 – 303 Mm³ annually), indicating superior performance of forecast-informed flexible operations. As Figure 13c illustrates, hydropower generation is maximized during high flow months (July – December). A systematic increase in the minimum yearly and monthly energy production and water allocation is critically important from users’ perspectives, as the value of improving upon unfavorable conditions often outweighs that of more acceptable conditions (Tversky and Kahneman 1992).

Comparison of scenarios during climate extremes is also insightful for evaluating the benefit of reservoir operational strategies for the end-user. For instance, under current operations (Scenario A), energy production and water allocation are low in dry years and higher in wet years, as may be expected (391 GWh and 146 Mm³ compared with 624 GWh and 303 Mm³; Table 5). Although energy production and allocation are moderate during the dry years, the year directly following shows greatly decreased production and allocation (e.g. 237 GWh and 40 Mm³ for 1984; Figure 15). A similar trend follows for Scenario B, yet additional allocation allows a slight increase in post-dry year benefits.

**Table 5.** Prediction-coupled reservoir model benefits to the end-user, all scenarios during a dry year (1983), post-dry year (1984) and wet year (1991).

<table>
<thead>
<tr>
<th>Model Scenario</th>
<th>Reservoirs Modeled</th>
<th>Forecast</th>
<th>Total Energy (GWh)</th>
<th>Total Water Allocation (Mm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dry</td>
<td>Post-Dry</td>
</tr>
<tr>
<td>A. Current Operations</td>
<td>Finchaa</td>
<td>--</td>
<td>391</td>
<td>237</td>
</tr>
<tr>
<td>B. Meet Agriculture Demand</td>
<td>Finchaa</td>
<td>--</td>
<td>252</td>
<td>320</td>
</tr>
<tr>
<td>C. Hydropower Maximizing</td>
<td>Finchaa</td>
<td>STAT</td>
<td>365</td>
<td>499</td>
</tr>
</tbody>
</table>
D. Hydropower Maximizing and Amerti

| STAT | 508 | 552 | 952 | 227 | 261 | 531 |

Note: AG - agriculture, HYD - hydropower, AVG - average of both users, STAT - statistical, seasonal forecast.

Figure 15. Annual energy generation (a) and water allocation (b) for the Finchaa reservoir for current operations (Scenario A), meeting agricultural demand (Scenario B), and maximizing hydropower (Scenario C) during dry (1983), post-dry (1984) and wet (1991) years, respectively.

The benefits of Scenario C are particularly evident during extreme conditions. While energy production and water allocation are at a minimum during very dry conditions, the ability to foresee drought allows the operational policy to shift such that energy production and water allocation the year following do not drop considerably (Table 5, Figure 15), as is the case in Scenario A. Granted, in this case, the following year is a near-average year, however a similar outcome may also be expected for two subsequent dry years, with the second year faring better when a flexible rule based on predicted climate information is included. The challenge for Scenario A – conditioned on a static rule curve – is that although it may perform well in the dry year in terms of energy production and agriculture allocation, as significant volumes of water are passed through the turbines, this results in a severe depletion of the reservoir and an inability to similarly provide in the following year. Outcomes from Scenario C for dry years and the year following – in which the
effects of the one dry year are spread across multiple years – is also seen as favorable by users who typically desire minimizing inter-annual variability in energy production and agricultural allocation. Comparing across Scenarios A, B, and C for a dry year, although benefits do not necessarily always meet agricultural demands (shown by deficits in Figure 13c, d), Scenario C performs best based on examination of total benefits.

_Oppportunity for Future Re-operation of the Finchaa – Amerti System_

Scenario C utilizes all available water in the Finchaa reservoir in numerous years, and is water limited, not energy production limited (turbine size). Thus, to expand energy production and potentially better satisfy all downstream agriculture demands, additional supply is needed. In the last couple decades, a reservoir was created on the adjacent Amerti River to serve as additional storage for the Finchaa system. Scenario C is therefore again considered but for the multiple reservoir (Finchaa – Amerti) system, named Scenario D.

Not surprisingly, the addition of the Amerti reservoir increases overall performance. High reliability (0.64 average; 88% percent increase over Scenario C) and resilience (0.24 average; 75% increase) as well as low vulnerability (2.30 average; 42% increase) illustrate the benefit of access to additional water supply (Table 4). Energy generation (494 – 990 GWh yearly; Table 4) and water allocation (221 – 556 Mm$^3$ yearly) are the highest across all scenarios, particularly in regard to minimum monthly and annual production/allocation, most often associated with dry years and immediately following (Table 5; Figure 13; Figure 16). While Scenario D is clearly superior at meeting sectoral demands in most years (Figure 16), there are years where Scenario B outperforms. For example, Scenario D adopts releases in anticipation of drier years in 1994 – 1996, and thus performs less well compared with Scenario B. However, given the static Scenario B releases full agricultural demands, performance suffers in 1998 due to depleted reservoir volumes, with performance still below that of Scenario D for the few years following. The Finchaa reservoir alone does not have the inflow capacity to satisfy agriculture demands (26.3% of years demand is met), and while the Finchaa-Amerti system performs better (63.2% of years demand is met), 100-percent reliability is still not achievable considering the historical record (Figure 17).
Figure 16. Annual downstream water allocation under varying scenarios for the Finchaa (solid) and Finchaa-Amerti (dashed) systems.

Figure 17. Annual inflow volume to the Finchaa, Amerti, and combined reservoirs. Annual agricultural water demand denoted by red line.
CHAPTER 4: DISCUSSION AND CONCLUSIONS

The BNB in Ethiopia provides critical natural resources for the nation and continent, yet spatial and temporal variabilities in climate adversely affect agriculture, energy, and other sectors. Advanced predictions of seasonal moisture to the region during the dominant JJAS season may provide information to enhance water resource decisions, yet the disconnect between the spatial scale upon which skillful predictions are issued and the sectoral decision-making scale renders current predictive information inadequate in many cases. Through development of skillful local-scale statistical, seasonal precipitation and streamflow predictions, coupled with a multi-purpose reservoir model, this work offers an alternative means to bridge the existing disconnect between prediction and decision-making scales for benefit to agriculture and energy sectors in Ethiopia.

Chapter two focuses on evaluating local-scale precipitation predictions to potentially enhance agricultural decision-making in a Blue Nile basin location using statistically downscaled NMME dynamic models, a regional cluster-based model, and a local statistical model. Statistical downscaling of dynamic models and the regional cluster-based model result in low prediction skill at the local scale, as compared to the local statistical model predictions. While the use of dynamical models is potentially appealing from an operational perspective, the modest skill and model complexity are clear drawbacks. Statistical methods have been utilized for decades in seasonal climate prediction, and while not a novel approach, results here are clearly superior, showing promise for enhanced application to water resources management.

It is generally unsurprising that statistical methods conditioned on local scale teleconnections can result in higher prediction skill at the local scale. One clear advantage of statistical approaches is the ability to condition relationships of interest at the local scale using the historical record without the need to fully capture complex climate dynamics at the global or regional scale. In a region with many teleconnections to global climate phenomena, such as ENSO and the Indian Monsoon, statistical methods may be better positioned to capture smaller-scale inter-annual factors that modulate variability. For the BNB, Ethiopia, regional variability in precipitation and streamflow can differ markedly from local variability. For example, correlation of Koga and Finchaa JJAS precipitation with BNB region-wide precipitation are only 0.51 and 0.70, respectively. These
differences illustrate the importance of capturing heterogeneous local climate patterns that diversely control moisture transport to relatively proximal locations. While global climate predictors may be similar regionally, there may be important nuances. Statistical approaches have demonstrated hydroclimatic prediction skill in a variety of settings (e.g. snow-dominated basins: Delorit et al. 2017, drought-prone regions: Mortensen et al. 2018), suggesting transferability of methods outlined here may be promising, however dominant, identifiable wet seasons (or high flow seasons) and clear climate signals are still likely necessary.

It is important to note that alternative statistical or dynamic bias-correction approaches could be evaluated, including linear interpolation, spatial disaggregation, or downscaling with regional climate models, which may increase local scale prediction skill (Fowler et al. 2007; Maraun 2013; Wood et al. 2004). However, this may only serve to increase complex resource requirements for a limited return. Likewise, statistical modelling methods have limitations, such as short (37 years in this case) time series of observational data that are potentially very sensitive to single anomalous years. For example, to compare with drop-one-year cross-validation, the 37-year historical record was split into a 19-year period for model calibration and 18-year period for model validation by random selection (and the process repeated hundreds of times). For a June 1 forecast model, the median correlation is 0.44 for the calibration period and only 0.07 for the validation period. This drop in correlation value is not altogether surprising considering the limited number of years in the time series but does highlight the potential sensitivity of models calibrated on a short observational record. One avenue for future work is to explore the Community Earth System Model (CESM; Kay et al. 2015; Ridge et al. 2013) with thousands of representative years to extend the “observational” record and potentially better characterize modeling uncertainties.

Model selection is intricately linked with spatial and temporal prediction scale. Recognizing the limitations of the current study, results indicate that both dynamic and statistical models have value for hydroclimatic prediction, with the decision of model structure based on the scale at which prediction information is required. Although many studies demonstrate value in predictions at the local scale, few local-scale operational prediction models exist. This work shows promise for bridging this gap through empirical forecast methods in Ethiopia.
One manner in which the value of hydroclimatic prediction models and information can be assessed is by the degree to which predictions can aid water resource planning and management decisions, particularly in preparing for extreme conditions. Multiple factors, including model structure or complexity, prediction skill, scale, lead time, and uncertainty evaluated here do play an important role in determining prediction value. As demonstrated, at larger (basin) spatial scales, the NMME dynamic model structures outperform statistical prediction models based on both deterministic and categorical metrics. Yet, for decisions at the local scale, statistical model structures conditioned on oceanic and atmospheric climate variables better capture small-scale variabilities in moisture transport to the region, exhibiting greater prediction skill for predictions issued at the start of the season (June) and extending up to five months prior (January). Thus, careful consideration of model structure based on the targeted decision scale has the potential to increase the value of predictions to the end-user.

Further, Chapter three examines how local-scale hydroclimate predictions may be coupled with reservoir models to translate predictive information into actionable strategies that may aid water resource management decisions. Reservoir systems are a critical resource that support energy, agriculture and other sectors, and may allow communities to mitigate risks posed by climate variability. Yet, the management of reservoir systems is complex. In many cases, operational rule curves have been developed to guide management of the system. Significant climate variability can challenge static rule curve approaches. Two examples of integrating local-scale predictive information with reservoir management models are evaluated in this work: 1) using JJAS precipitation predictions to predict October irrigation reservoir volume, and 2) coupling local-scale streamflow predictions with a multi-purpose reservoir model to evaluate alternative operational strategies.

Precipitation predictions applied to a simple reservoir model for a June 1 prediction of end of season (October 1) Koga reservoir volume would likely have correctly predicted whether or not the reservoir would fill in 73% of years based on median predicted values, supporting farmers’ decisions regarding seed allocation and land preparation. Interaction with collaborators suggests that early indication of October 1 reservoir volume, while simple, has great value. For example, in 2015 the region experienced one of the worst droughts in the historical record and the Koga
reservoir was half-full following the JJAS season. Although farmers were told to prepare only half their land in expectation of decreased water allocation, this information was received too close to the planting season to inform decisions; thus, huge land preparation and labor costs were lost. Although skillful predictions do not necessitate enhanced decision-making on their own, they can actively contribute to the suite of information available to help communities partially mitigate risks associated with climate variability. Early indication in years where alternative decisions may be warranted is imperative to allow for adaptive decision-making.

The coupled prediction-reservoir model of the Finchaa – Amerti system also exhibits skill in maximizing energy production and meeting agricultural demands. The addition of new hydropower facilities and agricultural expansion in the basin has warranted a re-evaluation of system-wide operations, which this work begins to address. The value of flexible operational strategies with predictive information is especially apparent in the dry years and following by prescribing a reserved quantity of water in the dry year (more conservative allocation) to allow for more available water in the subsequent year.

Chapter three examines a few alternative operational strategies, but clearly not an exhaustive list. Expected energy and crop prices could inform additional optimization objectives and further quantify the benefits and tradeoffs of different sectoral allocation quantities and timing and operational strategies. Different energy production strategies, such as maximizing firm energy, may also be identified. While the integrated models illustrate potential for increased hydropower benefits on the Finchaa-Amerti system, this value can only be realized if there is a market for additional power generation. Further, a more rigorous assessment of prediction uncertainty and its translation through the reservoir model, may be warranted to better quantify the robustness of results across different patterns of climate variability.

If the motivation for climate risk management is to support community adaption to climate variability, then factors affecting forecast uptake – beyond forecast skill – must also be considered. Social science and communication literature has repeatedly suggested that climate forecast advancements are not readily integrated into local-scale decision making structures or utilized to their full potential by the intended audiences (Gilles and Valdivia 2009; Pennesi 2007, 2011; Plotz
et al. 2017; Ziervogel et al. 2010), however this is not without its challenges, including prediction uncertainty, scale mismatch between climate and prediction models, institutional or political constraints, and risk aversion (Gilles and Valdivia 2009; Pennesi 2007; Ziervogel et al. 2010). Communication and dissemination strategies for increased understanding and uptake suggest that strategies tailored to the local region are key (Gilles and Valdivia 2009; Pennesi 2007, 2013; Ziervogel et al. 2010). Thus, developing skillful forecasts at the local, decision-making scale, and integrating with sectoral models, as demonstrated here, is imperative to resolve the disconnect between the scale at which existing models are skillful and the scale at which prediction information is valuable to end-users. This can serve to provide actionable information to aid water resources planning and management decisions and reduce the vulnerability of communities to climate variability.
REFERENCES


seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction.”


