

The role of seasonal and subseasonal hydroclimatic prediction in disaster risk reduction: Applications to rainfed agriculture and extreme rainfall

by

Jonathan M. Lala

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
2022

Date of final oral examination: 01/20/2022

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

Paul Block (Advisor), Associate Professor, Civil and Environmental Engineering

Daniel Wright, Assistant Professor, Civil and Environmental Engineering

Daniel Vimont, Professor, Atmospheric and Oceanic Sciences

Justin Sydnor, Professor, Risk and Insurance

Liangzhi You, Senior Research Fellow, International Food Policy Research Institute

Dedication

those affected by disasters: the hungry, the thirsty, the stranger, and the sick

Acknowledgements

To my friends, family, coworkers, and mentors, I extend my greatest gratitude. To my advisor, Paul Block, I have truly enjoyed your mentorship and support. Without it, I would not be standing here today, defending a PhD dissertation after several years. You have managed to walk the fine line of providing guidance and support while allowing me to pursue my passions, and I can only hope that my future bosses will demonstrate this to the fraction of the degree that you have. To the rest of my committee, Dr. Daniel Wright, Dr. Daniel Vimont, Dr. Justin Sydnor, and Dr. Laingzhi You, thank you for your support, especially when I had to dive into specialist material outside my own field. To my coauthors—Dr. Seifu Tilahun, Dr. Meijian Yang, Dr. Guiling Wang, Dr. Juan Bazo, Vaibhav Anand, Mekdim Dereje, Dr. Ying Zhang, Dr. Liangzhi You, and Dr. Donghoon Lee—thank you for all your help in making our papers truly interdisciplinary. Finally, to my family, I thank them for their support throughout not just my PhD but also my life, granting me the stability to undertake a large project mostly on my own (although not without the tremendous help of those listed above). I especially thank my dad for having a mobile job, allowing me to see the world and develop the level of independence needed for my work and my life.

This research has been supported by the National Science Foundation under Grant No. 1545874, a UW2020 grant from the Wisconsin Alumni Research Foundation at the University of Wisconsin-Madison, and an NSF INTERN award associated with Grant No. 1845783.

Table of Contents

Dedication	i
Acknowledgements	ii
Table of Contents	iii
List of Figures	vii
List of Tables	xii
Abstract	xiii
Chapter 1. Introduction	1
1.1 Ethiopia	3
1.2 Peru.....	5
Chapter 2. Predicting rainy season onset in the Ethiopian Highlands for agricultural planning	7
2.1 Introduction	7
2.2 Data and methods	8
2.2.1 Onset definition	9
2.2.2 Climate signals	12
2.2.3 Prediction models	13
2.3 Results	16
2.3.1 Climate signal results.....	19
2.3.2 Quantitative forecast results	20
2.3.3 Qualitative forecast results	22

2.3.4 Application to planting dates	25
2.4 Conclusions and Discussion.....	27
Chapter 3. The role of rainy season onset uncertainty on maize yields in Ethiopia	32
3.1 Introduction	32
3.2 Methodology	35
3.2.1 Defining onset and planting dates	35
3.2.2 Modeling maize production.....	38
3.3 Results and Discussion.....	39
3.4 Conclusion.....	45
Chapter 4. Incorporating seasonality into an agro-economic model of Ethiopia	47
4.1 Introduction	47
4.2 Methodology	48
4.2.1 Baseline model	49
4.2.2 Seasonal input data and disaggregation.....	49
4.2.3 Incorporation seasonal climate variability.....	50
4.2.4 Forecast-informed yield gains for maize	51
4.3 Results	53
4.4 Discussion	57
4.5 Conclusion.....	59
Chapter 5. Optimizing forecast-based actions for extreme rainfall events in Peru	61

5.1 Introduction	61
5.2 The Peruvian Red Cross early action protocol for extreme rainfall.....	62
5.3 Modeling framework.....	66
5.3.1 Forecasts	67
5.3.2 Sensitivity analysis and trigger thresholds	69
5.4 Results	72
5.4.1 Forecasts	72
5.4.2 Sensitivity analysis and trigger thresholds	74
5.5 Discussion and Conclusions.....	79
Chapter 6. Evaluating prospects for subseasonal-to-seasonal forecast-based anticipatory action from a global perspective.....	85
6.1 Introduction	85
6.2 Methodology	87
6.2.1 Evaluation of disaster risk	87
6.2.2 Evaluation of forecasts	90
6.2.3 Identifying suitability scores	93
6.3 Results	93
6.3.1 Forecasts	93
6.3.2 Risk and suitability	94
6.4 Discussion	99

6.5 Conclusion.....	101
Chapter 7. Summary and recommendations	103
7.1 How can forecasts and climate information at sub-seasonal to seasonal scales be tailored to provide value to end users?	106
7.2 What are the strengths and weaknesses of various forecast valuation techniques?	107
7.3 How do different institutions and socioeconomic conditions affect the use and value of forecasts in disaster contexts?	107
7.4 Final words.....	108
References.....	109

List of Figures

Figure 1-1. Mean total precipitation (mm) for the <i>Belg</i> (February-May, left) and <i>Meher</i> (June-September, right) seasons in Ethiopia.....	4
Figure 2-1. Overview of study area (red box), showing mean onset date according to the yearly definition	10
Figure 2-2. Daily precipitation (left axis) and cumulative precipitation anomaly (right axis) for an example year, showing long-term yearly mean (green horizontal line, 4.1 mm d-1), long-term window mean (red horizontal line, 6.0 mm d-1), corresponding cumulative anomalies (measured by right axis), and onsets for each definition.....	11
Figure 2-3. Mean onset dates (left) and standard deviations (right) for the threshold (top), yearly (middle), and window (bottom) definitions in northeastern Ethiopia, 1981-2019, showing Koga study area (red box)	18
Figure 2-4. Historical onset dates and trends for three definitions	18
Figure 2-5. Ensemble hindcasts, threshold definition, for March 15 (leftmost bar) to May 1 (rightmost bar) issue dates, along with observed onset date (black line and circles) and climatological mean onset date (blue line)	21
Figure 2-6. Ensemble hindcasts, yearly definition, for March 15 (leftmost bar) to May 1 (rightmost bar) issue dates, along with observed onset date (black line and circles) and climatological mean onset date (blue line)	22
Figure 2-7. Ensemble hindcast, window definition, for March 15 (left bar) and April 1 (right bar) issue dates, along with observed onset date (black line and circles) and climatological mean onset date (blue line)	22

Figure 2-8. Qualitative hindcasts, threshold definition, using PLS regression (left) and random forest classification (right) for March 15 (topmost value for each square) to May 1 (bottommost value) issue dates. Values indicate percentage of years for which hindcasts fell in each category; a perfect hindcast would contain 33% for each value in the diagonal (green) boxes, and 0% in all other boxes	23
Figure 2-9. Qualitative hindcasts, yearly definition, using PLS regression (left) and random forest classification (right) for March 15 (topmost value for each square) to May 1 (bottommost value) issue dates	24
Figure 2-10. Qualitative hindcasts, window definition, using PLS regression (left) and random forest classification (right) for March 15 (top value for each square) and April 1 (bottom value) issue dates	24
Figure 2-11. Historical planting dates for each criterion	25
Figure 3-1. a) Elevation (m.a.s.l.), b) mean <i>Kiremt</i> (June-September) precipitation (mm), c) coefficient of variation of annual <i>Kiremt</i> (June-September) precipitation, and d) AEZs of Ethiopia	34
Figure 3-2. Mean yields (kg ha^{-1}) by woreda, 1979-2014, for the baseline (left), onset-informed planting (center), and forecast-informed planting (average of all simulations; right) scenarios ..	40
Figure 3-3. Mean observed yield, 2004-2013 (upper left), and relative biases of mean yields relative to observed yields, %, for the baseline (upper right), onset-informed (lower left), and forecast-informed (lower right) scenarios	41
Figure 3-4. Improvement in mean yields for onset-informed planting over baseline (%), left), forecast-informed planting over baseline (%), center), and onset-informed over forecast-informed planting (kg ha^{-1} , right)	41

Figure 3-5. Example yields, 1979-2014, from two woredas marked in Figure 3-4: (A) semi-arid climate, and (B) tropical highland humid climate, showing spatial heterogeneity in onset planting benefits	42
Figure 3-6. Improvement in expected utility for onset-informed planting over baseline (%), left), forecast-informed planting over baseline (%), center), and onset-informed over forecast-informed planting (kg ha^{-1} equivalent, right).....	42
Figure 3-7. Improvement in average value at risk, $\alpha = 10\%$, for onset-informed planting over baseline (%), left), forecast-informed planting over baseline (%), center), and onset-informed over forecast-informed planting (kg ha^{-1} , right).....	43
Figure 4-1. Gains in mean yield (left) and expected utility (right) over baseline conditions resulting from forecast adoption (Lala et al., 2021a).....	52
Figure 4-2. Cultivated area-weighted CYFs for <i>Belg</i> (left) and <i>Meher</i> (right) for five major cereal crops (note the difference in scales).....	53
Figure 4-3. National maize producer price for baseline (left) and forecast-informed approach (right)	54
Figure 4-4. GDP (left) and daily calorie consumption per capita (right) for baseline (top), and % gains by using forecast-informed strategy (bottom)	55
Figure 4-5. Rural (left) and urban (right) poverty rate for baseline (top), and change forecast-informed strategy (bottom)	56
Figure 4-6. Zonal changes in GDP per capita, % difference over baseline, averaged over 1981-2020, for <i>Belg</i> (left), <i>Meher</i> (center), and annual (right) seasons.....	57
Figure 5-1. Process flow chart of EAP forecast triggers	64

Figure 5-2. Map of El Niño-induced extreme rain hazard by district (Peruvian Red Cross, 2018; based on the methodology of Marin-Ferrer et al., 2017), showing Piura-Lambayeque study region (black box)	66
Figure 5-3. Monthly precipitation (PISCO; Aybar et al., 2019) in the Piura-Lambayeque study region	66
Figure 5-4. Observed monthly precipitation (black line) and NMME modeled precipitation (colored lines) from 1995-2000, showing extreme El Niño event of 1997-98.....	73
Figure 5-5. Brier skill scores for monthly and seasonal hindcasts over 1982-2018	74
Figure 5-6. Disaggregated Brier scores (i.e., $(pt - ot)^2$) for the monthly and seasonal PLS forecasts (top) and weekly forecasts (bottom)	74
Figure 5-7. Range of possible benefit-cost ratios per season for probabilistic trigger thresholds of 0.1-0.99 and C/L_a ratio of 0.1 for PLS regression (blue boxes) and logistic regression (black boxes). Maximum possible flood-related losses (based on the number of extreme events) illustrated with a black line (right axis)	75
Figure 5-8. Benefit-cost ratio (top) and relative expense reduction (bottom) averaged over 1982-2018 for various probabilistic trigger thresholds and a C/L_a ratio of 0.1, using PLS forecasts (left) and logistic regression forecasts (right)	77
Figure 5-9. Utility with $A = 0.01$ (top) and $A = 0.5$ (bottom), for a C/L_a ratio of 0.05 (left) and 0.5 (right), using logistic regression forecasts	77
Figure 5-10. Combined false alarm ratio averaged over 1982-2018 for various probabilistic trigger thresholds, using PLS forecasts (left) and logistic regression forecasts (right).....	78
Figure 5-11. Maximum benefit-cost ratio as a function of C/L_a ratio for the two forecast methods, a perfect forecast (the inverse of the C/L_a ratio) and a feasibility threshold ($B/C = 1$)	79

Figure 6-1. Components of INFORM risk index (source: INFORM, 2020)	88
Figure 6-2. Peak month of precipitation, based on highest mean, 1981-2020.....	91
Figure 6-3. Maximum Brier skill scores for NCEP and ECMWF forecasts (BSS_{max}), for peak month (left) and season (right), and for the 15 th percentile (top; droughts) and 95 th percentile (bottom; floods) of monthly precipitation	94
Figure 6-4. Country-level risk score (INFORM, 2020; top), suitability score for peak month (middle) and peak season (bottom), for droughts (left) and floods (right)	95
Figure 6-5. Greater Horn of Africa: risk score (INFORM, 2020; top), suitability score for peak month (middle) and peak season (bottom), for droughts (left) and floods (right)	97
Figure 6-6. Peru: risk score (top), suitability score for peak month (middle) and peak season (bottom), for droughts (left) and floods (right).....	99

List of Tables

Table 2-1. Pearson's correlation coefficient (above diagonal) and mean absolute difference (days, below diagonal) between onset dates for each definition	19
Table 2-2. Predictor variables, correlation with onset, and PLS coefficients	19
Table 2-3. Average prediction error (days) and rank probability skill scores (RPSS) for each forecast (by onset definition and issue date) using PLS regression, including percent reduction in error over climatology (prediction of mean onset date) and percent of onsets occurring before issue date	20
Table 2-4. Number of historical false onsets based on naïve planting criteria and number of false onsets in same years when planting according to forecasts for various issue dates	26
Table 3-1. Share of area (%) that demonstrates improvement over baseline by planting strategy and performance metric	43
Table 5-1. Costs and benefits of weekly interventions as a ratio of maximum achievable benefits <i>B</i> and full cost of intervention <i>C</i> . Note that unavoidable losses are not included, as there may be more than one loss event per season.....	70

Abstract

Globally, water resources are under increasing stress and exacerbated by a variable climate. In low- and middle-income countries where lives and property are most exposed, climate shocks can have dire consequences, leading to poverty traps at the individual level and significant losses in economic growth at the national level. The proliferation of climate forecasts, particularly at the seasonal and subseasonal scale, offer some promise in mitigating risks from natural disasters, but they must be adequately adapted to local decision-making to be of full value. Moreover, this value is often inadequately assessed due to model assumptions and generalizations that only capture a part of the end users' true valuation, along with a common lack of co-evaluation with end users. A variety of models, forecast lead times, valuation metrics, and case studies, contrastingly, provides a far more nuanced understanding of forecast and climate information value, particularly when co-evaluated with end users themselves.

This dissertation outlines the role of climate forecasts in disaster management decisions and provides an understanding of the value associated with climate forecasts for two case studies, in Ethiopia (Chapters 2-4), Peru (Chapter 5), and both (Chapter 6). Chapter 2 describes the development of a subseasonal statistical forecast to predict the onset of the rainy season, which may inform farmers by guiding planting times. Chapter 3 applies this forecast to maize yields across the country using a process-based crop model, demonstrating how optimal planting time can buffer against climate shocks. Chapter 4 incorporates seasonality into an economy-wide, partial equilibrium multimarket model of Ethiopia to capture broad-scale implications and interactions from the uptake of climate forecasts. Chapter 5 includes a sensitivity analysis of early action plans for extreme rainfall that combines forecasts at multiple lead times to optimize decision

variables for loss minimization. Chapter 6 combines global forecast skill with socioeconomic data on risk to identify regions where anticipatory action for drought and flood preparedness may be most suitable. These combinations of qualitative, quantitative, and economic valuation techniques, all developed in coordination with partners and end-users, provide insights into appropriate application of such forecasts to be of benefit to society.

Chapter 1. Introduction

Global water resources are under increasing stress and variability as populations and the number of extreme events increase, creating challenges for water managers (Ferguson et al., 2018). Worldwide, direct financial losses from natural disasters average over \$165 billion per year; smaller indirect losses raise that figure by at least 50% (World Bank, 2014). Both low- and middle-income countries bear the brunt of these disasters: a growing global middle class—numbering five billion by 2030—retains much of their wealth in property, and middle-income countries capture a growing share of disaster losses, accounting for nearly 3% of GDP. Likewise, the minimal physical assets the poor have are often highly exposed, subjecting them to poverty traps (Borgomeo et al., 2017). Physical infrastructure-based solutions have existed since the dawn of water and disaster management, but they often have unintended consequences that make non-physical management, such as conservation, preferable in some cases (Di Baldassarre et al., 2018). The emerging field of planetary health proposes one alternative approach, in which adaptation to climate is made through proper decision making without unsustainably impacting the natural environment (Horton and Lo, 2015). Such non-structural measures, moreover, tend to be more cost-efficient than their physical counterparts (Kelman, 2013). There is thus a growing need for increased operational capacity through increased preparedness, minimizing excessive investments in physical infrastructure, to respond to climate variability.

The benefits of forecast application in water resource operations have been illustrated extensively, with an emphasis on the need to tailor forecasts to specific sectors and locations (Alexander et al., 2018). Weather forecasts—which predict hydrometeorological conditions hours to days in advance—are well-developed and used worldwide; climate forecasts—issued weeks to years in

advance—are still less universally common, however their application is growing. Seizing upon this trend, the Subseasonal-to-Seasonal (S2S) Prediction Project (WMO, 2018) was launched in 2009 to improve climate predictions, particularly for extreme or high-impact events. A decade later, the second phase of the project was initiated with the goal of “enhancing operational infrastructures and user applications” (WMO, 2018). Real-time operations require the cooperation of on-the-ground organizations, such as government agencies, agricultural extension agencies, and non-governmental organizations (NGOs), that can properly connect scientific findings to end-user benefits. Forecasts can provide considerable support in this decision-making process, yet their value can be measured in various ways—for example, decision theory, economic models, or benefit transfer models—that alone may not be fully representative of end users. Continuing research in the field, then, must integrate various methodological approaches and incorporate co-evaluation with forecast users (Soares et al., 2018). This context leads to three framing questions addressed over the course of research summarized in this dissertation:

1. How can forecasts and climate information at sub-seasonal to seasonal scales be tailored to provide value to end users?
2. What are the strengths and weaknesses of various forecast valuation techniques?
3. How do different institutions and socioeconomic conditions affect the use and value of forecasts in disaster contexts?

The research covered in this dissertation explores the role of forecasts through a variety of value metrics, scales, and sectors, using case studies in two different regions. The first, in Ethiopia, describes the development of a subseasonal forecast to predict rainy season onset (Chapter 2), coupled with a process-based crop model (Chapter 3) and partial equilibrium model (Chapter 4), to identify potential local yield gains and regional or national economic effects. The second, in

coastal Peru, considers the operational capacity of an anticipatory action framework for extreme rainfall using forecasts at a variety of temporal and spatial scales (Chapter 5). Finally, this dissertation concludes with a global analysis of anticipatory action framework suitability for disaster risk reduction by considering both forecast skill and a multi-dimensional approach to risk (Chapter 6). Ethiopia and Peru are highlighted as case studies in this final, global analysis.

1.1 Ethiopia

Ethiopia's economy is primarily agricultural, with farming and husbandry employing 70% of the labor force and contributing to approximately 35% of GDP (CIA World Factbook, 2020). Farming is highly seasonal, aligning with two rainy seasons (Figure 1-1). Most of the country relies on a summer rainy season—the *Kiremt*—for 65-95% of annual rainfall; droughts and famines are often associated with a failure of these rains (Segele et al., 2015; Temam et al., 2019). There is also a shorter and more volatile spring rainy season—the *Belg*—mainly used for crop production in the southwest of the country, which is also subject to drought and often a precursor of wider *Kiremt* failure (Temam et al., 2019). Under current conditions in Ethiopia, the negative effects of climate variability outweigh positive gains (Block et al., 2008), which prompts decision makers to employ risk reduction strategies. Regions with highly seasonal rainfall are particularly ripe for forecast-informed climate services, with farmers often ranking climate variability as a major constraint (Rao et al., 2019). Studies in the drought-prone Tigray and Somali regions specifically found that early action measures could have saved humanitarian agencies \$1.2 billion USD over 15 years (Cabot Venton, 2018). In neighboring Kenya, effective use of weather information was found to potentially increase gross margins for maize by up to 70 percent for perfect information and up to 24 percent using sea surface temperature-based forecasts (Hansen et al., 2009). Moreover, although a changing climate is projected to result in increased total precipitation in the Nile Basin

across the calendar year, projections for the core months of the Kiremt are much less clear (Ferguson et al., 2018).

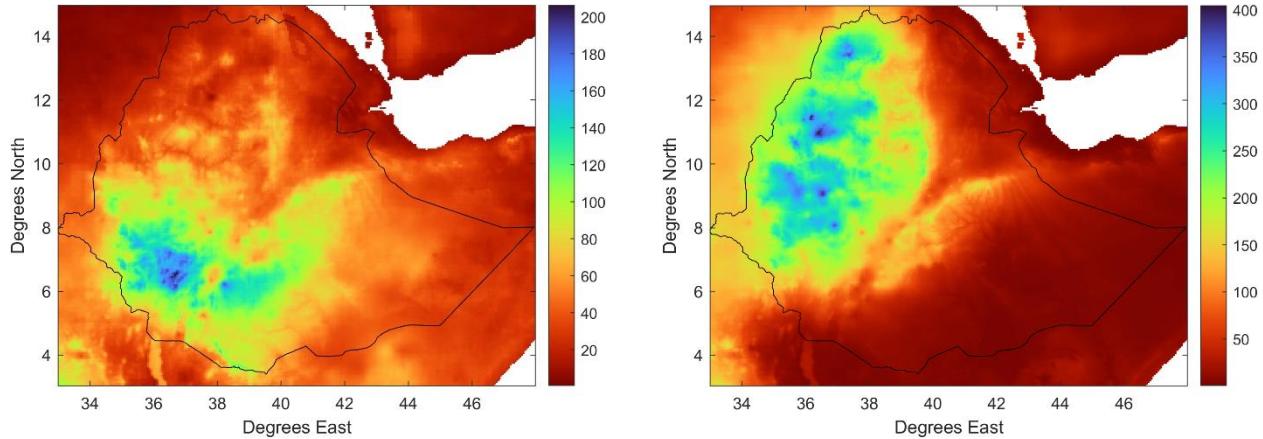


Figure 1-1. Mean total precipitation (mm) for the *Belg* (February-May, left) and *Meher* (June-September, right) seasons in Ethiopia

The hydroclimatology and seasonal prediction literature in Ethiopia is rich; however, relatively few studies consider intraseasonal characteristics such as the timing of the wet season. Those that have predicted the onset and cessation of rains—notably MacLeod (2018)—have mainly utilized dynamic models, which are of coarse spatial resolution and may not be fully compatible with local decision-making (Wambui, 2019). Contrastingly, statistical models may be well-suited to places with heterogeneous geography such as Ethiopia (Zhang et al., 2016; Alexander et al., 2019), and may better capture extreme events (Nicholson, 2017). Valuation of these models is most effective by using a variety of methods and active co-evaluation with forecast users (Soares et al., 2018). We thus propose the development of statistical forecast models and climate informatics in cooperation with end-users—predominantly farmers and NGOs—through an interdisciplinary project funded by the National Science Foundation. Three chapters of this dissertation consider different aspects of this approach: Chapter 2 demonstrates the development of a subseasonal forecast model of the onset of the rainy season and its application to maize planting, particularly

by reducing the risk of planting during a false onset. Chapter 3 extends this analysis by integrating forecast knowledge of onset into a process-based crop model, demonstrating potential gains associated with optimal planting times. Finally, Chapter 4 provides an assessment of the economic value of seasonal and subseasonal knowledge through a country-wide multimarket partial equilibrium model to investigate how local actions may reverberate through the national level economy and society.

1.2 Peru

Coastal Peru contains more than half of the country's population but is extremely arid and generally receives only sporadic rainfall, mainly in the austral summer (Rau et al., 2016). Extreme precipitation events, however, are not rare. These events and interannual variability are generally associated with the El Niño – Southern Oscillation (ENSO) phenomenon, with extreme El Niño years bringing potential catastrophic rainfall to the coast. The most recent, in 2017, affected half a million people in the Piura Region alone; damages were valued in the billions of USD even though extreme rainfall was mostly confined to a single month (INDECI, 2017). These extreme events, combined with the relative predictability of ENSO and support from institutions in Peru, have resulted in the development of early warning systems throughout the country. Simultaneously, there has been an increased interest in index-based (or parametric) insurance, which triggers payouts based on a pre-defined threshold in lieu of traditional claims adjustment. This has been particularly applicable in low- and middle-income countries, as it requires few observations, is standardized and transparent, and reduces administrative costs, moral hazard, and adverse selection (Skees, 2011). Coupling forecasts and the index insurance concept, a number of NGOs have piloted forecast-based financing (FbF) initiatives to take anticipatory action and disperse emergency funds and supplies prior to a disaster occurring. The potential costs of acting in vain,

however, must be explicitly considered. Risk to the disaster response unit is thus transferred from *ex-post* disaster recovery to *ex-ante* preparation. Forecast-based index insurance was implemented in northern Peru by 2010, while a forecast-based emergency response for floods pilot project began in 2015 (Coughlan de Perez et al., 2015; Lopez et al., 2017). The momentum for FbF programs is therefore growing both in Peru and worldwide.

Nonetheless, outstanding questions and challenges within the FbF framework remain. Lopez et al. (2017) evaluate an FbF pilot using cost reduction as an optimization measure; however, the framework is based on weekly forecasts, leaving little room for prepositioning, a potential option with seasonal or subseasonal forecasts. Bischiniotis et al. (2019) consider a two-stage FbF framework to establish optimal lead times but do not explicitly optimize for trigger thresholds. Moreover, nearly all studies rely on a cost-loss ratio framework that assumes risk neutrality by the decision maker, despite empirical evidence that most are risk-averse (Matte et al., 2017). Finally, although socioeconomic elements of risk are implicitly considered in most FbF programs, there seems to be no standardization of risk that can be applied globally, despite its strong potential for integration with early warning (OCHA, 2021). Several issues therefore remain, including but not limited to explicit optimization of trigger thresholds, incorporation of risk aversion as opposed to a linear cost-loss framework, the role of seasonal and subseasonal forecasts in prepositioning and training, and the impact of socioeconomic conditions in the affected community. The final two chapters explore these outstanding questions: Chapter 5 presents an optimized operational framework for extreme rainfall in coastal Peru using multiple types of forecasts and lead times, and Chapter 6 details a global analysis of anticipatory action suitability by considering socioeconomic components of risk and forecast skill.

Chapter 2. Predicting rainy season onset in the Ethiopian Highlands for agricultural planning

Adapted from: Lala, J., Tilahun, S., and Block, P. (2020). Predicting rainy season onset in the Ethiopian Highlands for agricultural planning. J. Hydrometeorology, 21, doi:10.1175/JHM-D-20-0058.1

2.1 Introduction

Ethiopia's main rainy season, the *Kiremt*, occurs during the boreal summer and is responsible for 65-95% of total annual rainfall in the country, making it the primary driver of agricultural production (Segele et al., 2015). Agricultural planning, livestock herding, and reservoir management all rely on these rains, largely affecting national welfare. The tragically reoccurring droughts that have plagued East Africa's most populous country for centuries are most often associated with a failure of the Kiremt rains (Lanckriet et al., 2014). Past disasters and future climate uncertainty have prompted the Ethiopian government to specifically identify early warning systems and development insurance as key strategies in the country's National Adaptation Plan (Government of Ethiopia, 2019). While progress in this area is evident (Drechsler and Soer, 2016; Ewbank et al., 2019), the need to identify additional tools to specifically tailor local decision-making is significant.

Year-to-year variability in rainy season precipitation is a major challenge for farmers in East Africa, yet proper management may bring about positive gains or avoided losses. Many studies analyze the hydroclimatology of East Africa, including at the seasonal scale (e.g. Hasenrath et al., 2011; Moron et al., 2013; Yang et al., 2015), but relatively less attention is paid to intra- or sub-seasonal characteristics, despite their contribution to seasonal variability (Berhane and Zaitchik,

2014; Nicholson, 2017). Conversations with our partners in the field highlight uncertain rainy season onset as a key concern to farmers, especially given its increasing variance in timing (Darabent et al., 2020) and its influence on yield and crop prices later in the year (Davenport et al., 2021). A skillful prediction of Kiremt onset, therefore, coupled with effective dissemination of information, may allow farmers to plant crops in a timely manner, guide pastoralists in their search for pasture, and facilitate strategic reservoir operations.

In contrast to seasonal total precipitation, onset definition is much less clear. Definitions range from arbitrary thresholds, to precipitation anomalies, to non-precipitation-based phenomena, such as the behavior of local flora and fauna or changes in atmospheric patterns. These framings, based on observations, can produce widely varying onset dates—from as early as April to as late as July—necessitating an analysis of their strengths and weaknesses, particularly in the context of water management and agricultural planning. This chapter presents development and analysis of statistical onset prediction models and their application to multiple onset definitions. Application of forecasts to early-season planting—particularly for maize—is also considered. The Koga watershed in the Ethiopian highlands is selected for demonstration due to its status as an intensive agricultural region and because of the presence of a major reservoir at its outlet.

2.2 Data and methods

The approach adopted here includes determining observed onset annually according to three definitions, correlating onset date with large-scale pre-season climate drivers that may serve as predictors, and producing season-ahead onset date predictions at four different lead times using quantitative and qualitative methods. Comparisons are then made with farmers' planting decision calendar to demonstrate the utility of an onset forecast.

2.2.1 Onset definition

Onset date is characterized according to three definitions: a threshold-based definition (*threshold definition*), a definition based on precipitation anomalies relative to the long-term average (*yearly definition*), and a definition based on precipitation anomalies relative to April-July long-term averages (*window definition*). Daily precipitation values from 1981-2019 are taken from the Climate Hazards Group InfraRed Precipitation with Station dataset (CHIRPS; Funk et al., 2014), based on 0.05° satellite imagery and corrected with in-situ gauge data; this dataset has been shown to demonstrate very low bias over northwestern Ethiopia (Dinku et al., 2018). To evaluate the potential predictive skill based on a statistical approach, a primarily rainfed agricultural area upstream of the Koga reservoir (11.05° to 11.35° N, 37.00° to 37.35° E) is considered. Onset date is calculated at each 0.05° grid independently across the area, with the median value from all grid cells for each year retained for forecast training and validation (Figure 2-1).

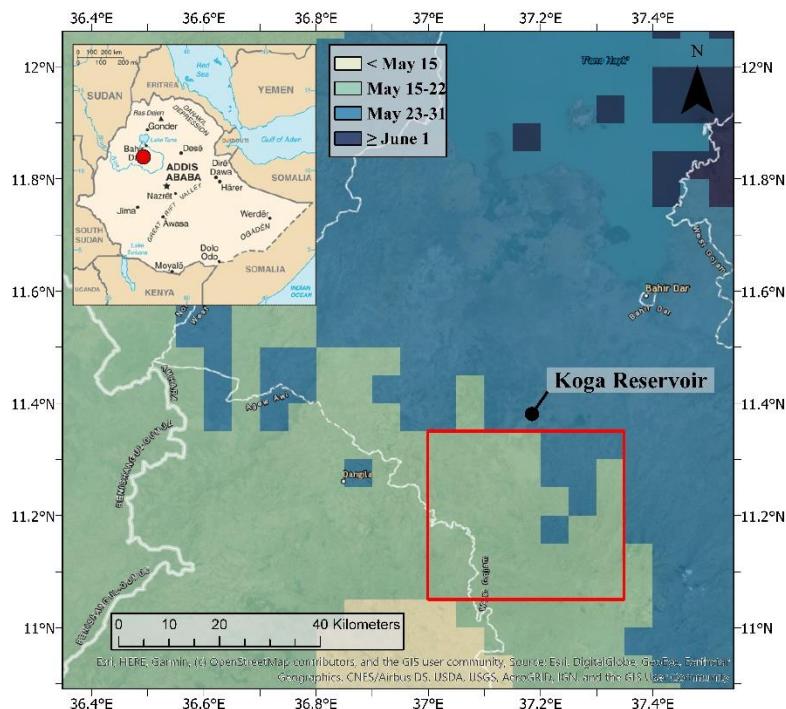


Figure 2-1. Overview of study area (red box), showing mean onset date according to the yearly definition

The *threshold definition* (Segele and Lamb 2005) is defined as a three-day accumulation of 20 mm or more after April 1, with no dry spells (< 0.1 mm each day for at least eight days) in the next 30 days. Although this is seemingly arbitrary, it is useful for local agronomy, acting as a proxy for soil moisture (MacLeod, 2018), and it was specifically developed for the relatively wet areas of northwest Ethiopia considered in this study. In contrast, two precipitation anomaly-based definitions are also considered. The *yearly definition* (Dunning et al. 2016) is calculated by taking the cumulative daily precipitation anomaly, starting January 1 of each year, relative to the long-term daily average. The global minimum of this cumulative time series for each year is defined as the onset date. As this definition makes no assumption of when the season occurs—in contrast to farmers, who are unlikely to plant outside of a given window regardless of precipitation conditions—a similar definition was adapted from MacLeod (2018) in which cumulative anomalies are calculated, starting April 1, relative to the April-July window only. This *window definition*, along with the yearly definition, is more likely to avoid false onsets (i.e. significant precipitation followed by a long dry spell), since the cumulative precipitation anomaly will be offset by the subsequent dry spell. Figure 2-2 illustrates each onset definition for an example year.

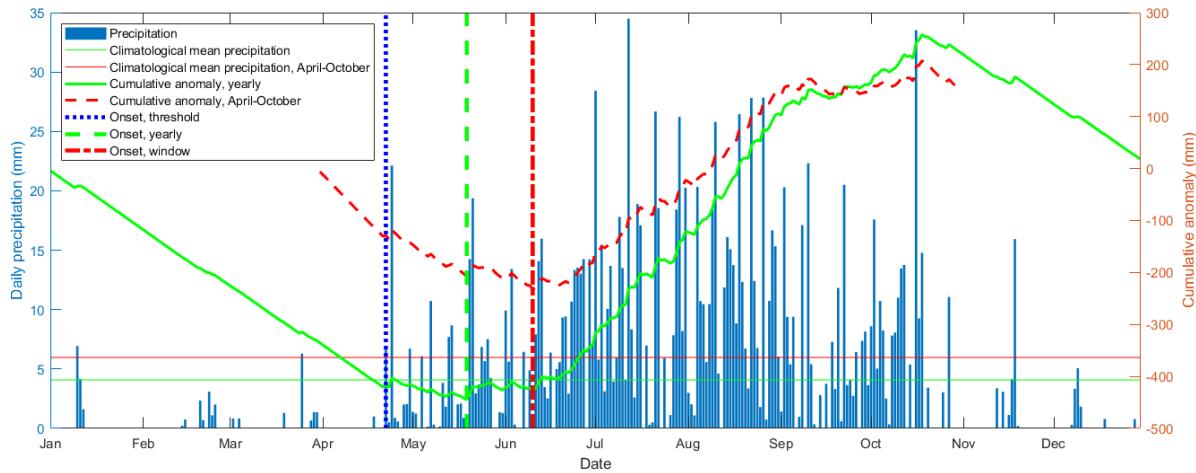


Figure 2-2. Daily precipitation (left axis) and cumulative precipitation anomaly (right axis) for an example year, showing long-term yearly mean (green horizontal line, 4.1 mm d⁻¹), long-term window mean (red horizontal line, 6.0 mm d⁻¹), corresponding cumulative anomalies (measured by right axis), and onsets for each definition

In addition to onset date, criteria for planting dates are also determined to test the utility of an onset forecast. Planting criteria typically rely on soil moisture, for which a threshold-type definition would be most apt (MacLeod, 2018). The threshold onset definition of Segele and Lamb (2005) is specifically designed for moist areas of Ethiopia that experience *Kiremt* rains (roughly north of 7° N), such as Koga, but farmers do not have perfect foresight about dry spells occurring in the 30 days after the rainfall threshold is reached. Maize is among the first crops planted in order to maximize growing season (Liben et al., 2015); however, farmers run the risk of planting during a “false onset”, in which a wet period is followed by a prolonged dry spell, which may reduce yield or require replanting (Kipkorir et al., 2007; Tadross et al., 2009). Discussions with farmers located within or near the Koga study area indicated a typical maize planting date around mid- to late May most years, which corresponds with mean onset date (Figure 2-1); however, dates varied widely, ranging from early May to early July in a survey of approximately 2500 farmers for the 2018

cropping season. A smaller survey of the 2019 cropping season found maize planting dates centered around early June for that year.

Considering the range of planting dates, the suitability of a threshold definition for planting purposes, and a lack of foresight regarding false onsets, three different planting date criteria are developed: (A) four days following a wet spell of 20 mm in three days on or after May 1; (B) four days following a wet spell of 20 mm in three days on or after May 15; and (C) four days following a wet spell of 50 mm in four days and at least one wet day (>0.1 mm) in the following three days on or after April 1. Criteria A and B are taken from the threshold onset definition (Segele and Lamb, 2005), but altered to eliminate foreknowledge of dry spells and to restrict planting dates to May or later, reflecting the farmer surveys. Criterion C is inspired by similar threshold methods in Tigray (Araya et al., 2012) and southeastern Africa (Tadross et al., 2009), but has slightly increased precipitation thresholds to account for the relatively wetter climate of the Koga region, as well as requiring an additional rainy day within the three days following a wet spell. Thus, the three criterion serve to reflect three sets of risk aversion, with farmers preferring to maximize growing season choosing Criterion A, risk-averse farmers choosing to wait for a longer spell of reliable rain choosing C, and the rest choosing B.

2.2.2 Climate signals

Historically, seasonal forecasts of East African precipitation have primarily relied on sea-surface temperature (SST) and sea-level pressure (SLP) from the Atlantic-Mediterranean, Indian, and Pacific Oceans, although recent research has also highlighted the value of atmospheric variables (Nicholson, 2017), some of which are considered in this study (e.g., geopotential height and zonal wind). Moisture transport to the Ethiopian Highlands is primarily sourced from the Atlantic Ocean, Indian Ocean, and Mediterranean Sea (Viste and Sorteberg, 2013; Segele et al., 2015). A trajectory

analysis conducted by Jury (2011) confirmed an Indian Ocean – Red Sea – Mediterranean Sea cyclonic behavior responsible for major floods in the region.

Selected oceanic and atmospheric variables (NCEP-DOE Reanalysis 2 dataset; Kanamitsu et al., 2002) from January – April (prior to onset) at a semimonthly timestep, that may serve as skillful predictors based on the literature described above, are correlated with onset date across 1981–2019. The month(s) with the highest correlation between onset date and potential predictor is retained. Further reduction in the number of predictor variables is achieved using the generalized cross-validation score (GCV; Craven and Wahba, 1979), balancing model error and the number of selected predictors, defined as

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

where e_t is the model residual, N is the number of data points, and m is the number of predictors. The set of predictors with the lowest GCV score for each onset definition and lead time is retained (see Table 2-2 for the full set of retained predictors).

2.2.3 Prediction models

Two model types are adopted for predicting onset: partial least squares (PLS) regression (Wold et al., 1984), producing a quantitative ensemble prediction of the exact onset date, and random forest classification (Breiman, 2001), providing a qualitative categorical (early/normal/late) prediction. Predictions are generated using both methods for four different issue dates—March 15, April 1, April 15, and May 1—to understand the trade-off between lead time and forecast accuracy. As an example, following Table 2-2, the March 15 issue date for the yearly definition uses four predictor

variables: (1) SLP averaged over the month of January and over the coordinates 35-40N, 0-10E; (2) 500 mb geopotential height averaged over the month of February and over the coordinates 10S-5N, 20-40E; (3) 1000 mb geopotential height averaged over March 1-14 and over the coordinates 35-40N, 0-10E; and (4) 1000 mb geopotential height averaged over March 1-14 and over the coordinates 10-40N, 20-45E. Note that for an April 1 issue date, predictors (3) and (4) are averaged over the entire month of March, and for an April 15 or May 1 issue date, a fifth predictor (500 mb zonal wind averaged over the coordinates 5-15N, 5-20E and over the April 1-14 or April 1-30, respectively) is included. Thus, for each predictor, there is a timeseries with a single value for each year from 1981-2019. This set of predictors is then used in the PLS regression or random forest classification. The prediction is still calculated in the case of an early onset for training purposes; hence “% onsets before issue date” is included as an independent skill metric (see Table 2-3).

PLS regression is calculated based on the z-scores of the selected climate predictors, unique for each onset definition. Predictors and responses are then decomposed according to:

$$X = TP^T \quad (2-2)$$

$$Y = UQ^T \quad (2-3)$$

where X is the set of predictors, Y is the set of responses, T and U are projections of X and Y , respectively, and P and Q are the loading matrices of X and Y , respectively (Ng, 2013). The decompositions are designed to maximize the covariance between X and Y , solving the multicollinearity problem, similar to principal component analysis, while also explicitly considering Y in the estimation of components (Abdi, 2003). Terms explaining more than 10% of the variance in the input data are retained, similar to Kaiser’s rule (Kaiser, 1960). Given the relatively small number of years ($n=39$) from which to fit the models, and given the statistical

independence in onsets between years, a hindcast across 1981-2019 is constructed in a drop-one cross-validation mode to evaluate performance. Ensemble predictions for each year are formed by bootstrapping, with replacement, from the residuals of the hindcast. Additionally, a qualitative prediction is generated from the deterministic model output, classifying as early, normal, or late onset based on the predicted date relative to the tercile of historical onset dates.

Separately, a random forest qualitative prediction is calculated using the same input data as the PLS regression and same classification criteria. Random forests are generated from a set of decision trees, with each tree generated by bagging (i.e., randomly selecting, without replacement, from examples in the training set); the general form is given as

$$\{h(\mathbf{x}, \Theta_k), k = 1, \dots\} \quad (2-4)$$

where h is the classifier function, \mathbf{x} is the set of inputs, and $\{\Theta_k\}$ is a set of independent but identically distributed random vectors. Classifications are made from the votes of each tree for the most popular class at input \mathbf{x} (Breiman, 2001).

Skill is measured using mean absolute error in days and rank probability skill score (RPSS; Wilks, 1995) for the quantitative forecasts and correct classification rate and extreme miss rates (i.e., forecast of early when observed late and vice-versa) for the qualitative forecasts. The RPSS provides a skill metric for ensemble forecasts by comparing the probabilities of categorical predictions with respect to a reference forecast, such as climatology. The rank probability score (RPS) is first calculated as:

$$RPS = \sum_{i=1}^n (F_{fct_i} - F_{obs_i})^2 \quad (2-5)$$

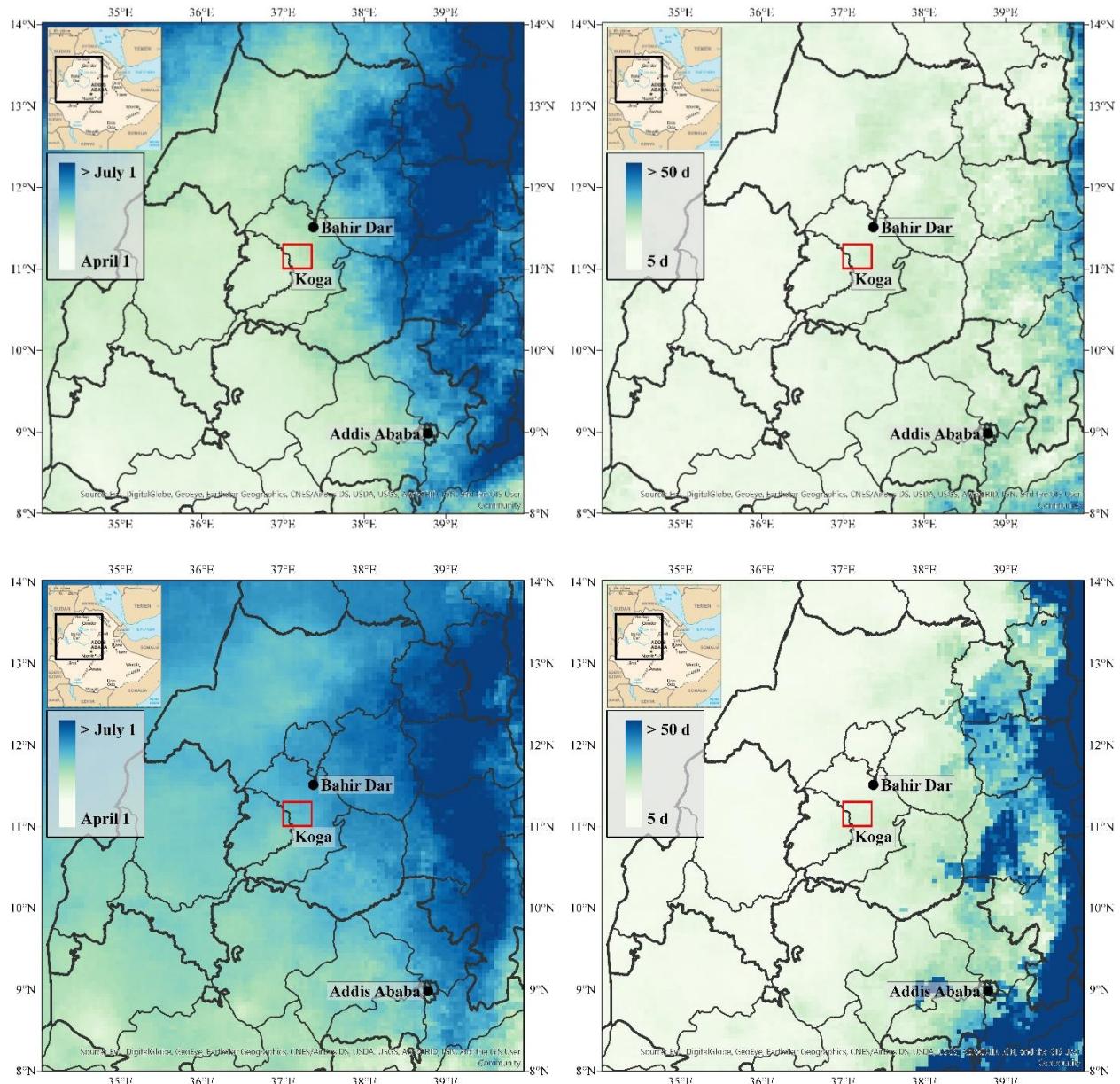
where F_{fct_i} and F_{obs_i} are the cumulative probabilities of the forecasts and observations, respectively, through category i , and n is the number of categories. The RPSS is then calculated as:

$$RPSS = \left(1 - \frac{RPS_{forecast}}{RPS_{climatology}} \right) \times 100\% \quad (2-6)$$

An RPSS of 100% indicates a perfect forecast, whereas scores below zero indicate inferior skill relative to climatology. For this study, the median RPSS over all hindcast years is taken as the relevant score. To capture statistically significant trends, the dates separating categories (early, normal, or late) are shifted each year by the slope of the trend line, uniquely for each onset definition.

2.3 Results

The *Kiremt* season generally observes a gradient of increasingly later onset as one moves northeast; this is true of all definitions of onset, although the anomaly methods tend to define onset later overall than the threshold method does (Figure 2-3). Variability in onset date follows a similar pattern among definitions, with eastern areas experiencing more variability than western areas. However, variability also dramatically increases at the southern and eastern edges of the *Kiremt* region, where a clear unimodal rainfall regime starts to transition into the bimodal regime of southern and eastern Ethiopia (Mengistu Tsidu, 2012).



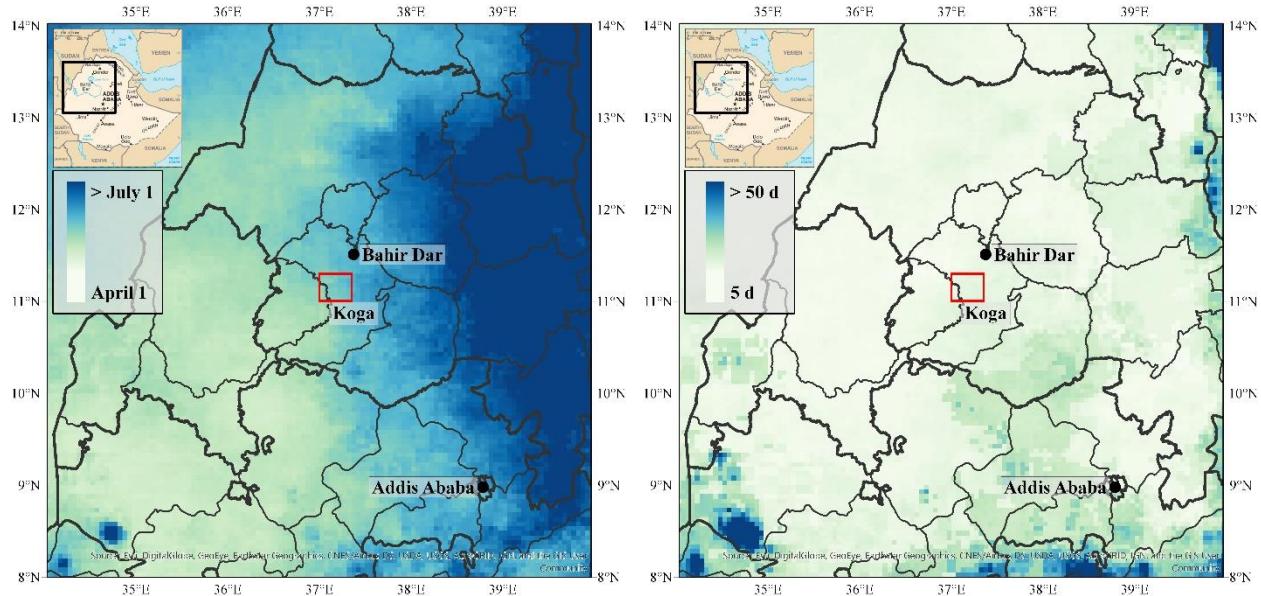


Figure 2-3. Mean onset dates (left) and standard deviations (right) for the threshold (top), yearly (middle), and window (bottom) definitions in northeastern Ethiopia, 1981-2019, showing Koga study area (red box)

For Koga, onset dates differed notably depending on the definition; early or late onset generally agrees between definitions however the specific dates vary widely (Figure 2-4, Table 2-1). Generally, the threshold definition produces the earliest onsets whereas the window definition estimates the latest onsets.

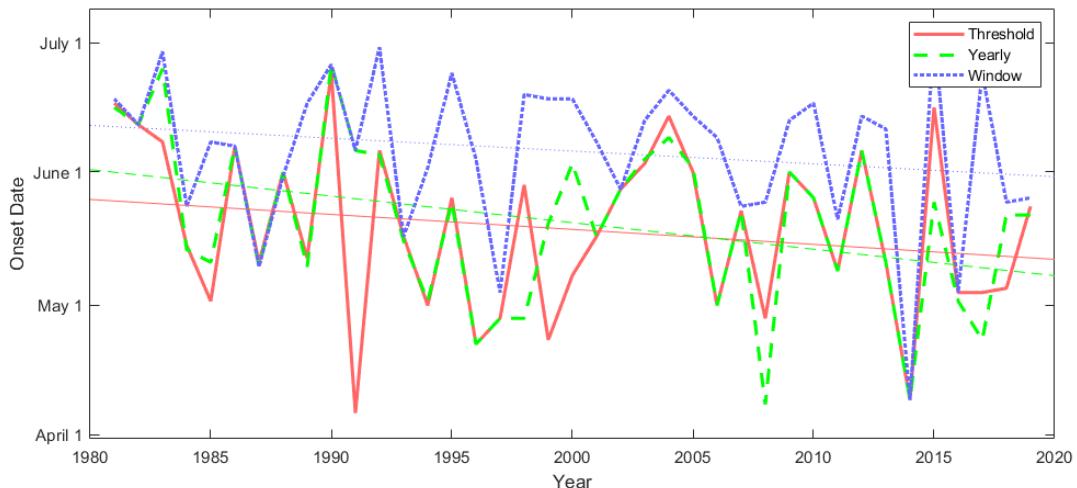


Figure 2-4. Historical onset dates and trends for three definitions

Table 2-1. Pearson's correlation coefficient (above diagonal) and mean absolute difference (days, below diagonal) between onset dates for each definition

	Threshold	Yearly	Window
Threshold	-	0.73	0.56
Yearly	6.7	-	0.56
Window	18.3	16.7	-

2.3.1 Climate signal results

In addition to SLP, which has been historically popular in seasonal forecasts, atmospheric variables—particularly geopotential height and zonal wind—are also found to be important contributors to onset across definitions (Table 2-2). Correlations with onset tend to be slightly stronger among the anomaly definitions, which have an intrinsic connection with total seasonal precipitation. In contrast, the threshold definition relies more strongly on specific weather events (e.g. a single period of heavy rainfall), which are more difficult to forecast given the long lead times considered in this study. Because of collinearity between predictors, correlations are not necessarily indicative of the strongest predictors, however; for example, 200 mb geopotential height over the Sahara has the lowest correlation coefficient with threshold onset, but the highest PLS coefficient.

Table 2-2. Predictor variables, correlation with onset, and PLS coefficients

	Variable	Region	Latitude	Longitude	Month(s)	Correlation	PLS coefficient			
							March 15	April 1	April 15	May 1
Threshold	SLP	Southern Ethiopia	0-5N	40-45E	Feb-Mar	0.43	6.46	3.58	2.05	1.62
		Southern Ethiopia	0-5N	40-45E	Mar	0.53	2.99	4.27	3.58	2.53
	Geopotential Height	Eastern Mediterranean (500 mb)	30-35N	25-35E	Feb	-0.42	-6.70	-5.03	-7.14	-7.99
		North Africa (1000 mb)	20-30N	30-40E	Mar	0.56	1.61	5.32	5.60	5.59
	Zonal Wind	Sahara (200 mb)	20-30N	10-30E	Apr	0.38	-	-	7.92	8.96
Window	Geopotential Height	Eastern Mediterranean (500 mb)	30-40N	25-35E	Jan-Feb	-0.57	-10.09	-9.12	-	-
		Western Pacific (1000 mb)	15S-5N	150-175E	Feb	0.39	6.26	5.20	-	-
	Zonal Wind	Central Sahara (1000 mb)	15-20N	0-10E	Mar	-0.52	-2.46	-5.62	-	-
Yearly	SLP	North Atlantic	35-40N	0-20W	Jan	0.45	8.14	7.44	6.07	6.15
		Eastern Mediterranean (500 mb)	30-45N	20-40E	Feb	-0.51	-7.82	-5.76	-5.98	-5.86
	Geopotential Height	Equatorial Pacific (1000 mb)	10S-5N	130-160W	Mar	-0.54	-	-6.71	-6.74	-6.73
		Mediterranean / Red Sea (1000 mb)	10-40N	20-45E	Mar	0.54	5.33	5.17	5.11	4.23
	Zonal Wind	Central Sahara (500 mb)	5-15N	5-20E	Apr	-0.41	-	-	-4.39	-5.18

SST is notably absent from all onset definitions, suggesting that surface forcings may play a lesser role in onset. Contrastingly, pressure variables (SLP and geopotential height) are featured as predictors in all onset definitions, with the Mediterranean and Red Sea featured as key regions. Predictors located in these regions exhibit high PLS coefficient values for the threshold and window methods, suggesting primacy in terms of influencing onset and confirming the findings of Jury (2011). For the yearly method, low-level pressures in the Atlantic and Pacific tend to dominate, although their influence is less stark relative to other variables, for which PLS coefficients are of similar magnitude.

2.3.2 Quantitative forecast results

PLS regression-based predictions demonstrate improvement for all definitions and all issue dates relative to climatology (Table 2-3). The GCV score for the window definition was lowest in the absence of April climate signals; hence, only a March 15 and April 1 forecast date are presented.

Table 2-3. Average prediction error (days) and rank probability skill scores (RPSS) for each forecast (by onset definition and issue date) using PLS regression, including percent reduction in error over climatology (prediction of mean onset date) and percent of onsets occurring before issue date

Definition	Average climatology error (d)	Issue date	Average prediction error (d)	% Reduction in error	% Onsets before issue date	Median RPSS	Median RPSS (early)	Median RPSS (late)
Threshold	16.4274	March 15	13.1017	20%	0%	31.56%	55.90%	40.64%
		April 1	12.3523	25%	0%	35.16%	79.19%	53.74%
		April 15	11.4669	30%	5%	49.38%	84.86%	54.55%
		May 1	10.7199	35%	23%	49.15%	86.72%	29.93%
Yearly	15.5464	March 15	12.0016	23%	0%	25.39%	68.25%	0.55%
		April 1	10.5222	32%	0%	40.37%	90.48%	36.41%
		April 15	10.0029	36%	5%	50.28%	95.95%	58.46%
		May 1	9.7161	38%	15%	54.98%	93.50%	69.74%
Window	13.6963	March 15	10.9614	20%	0%	14.23%	45.39%	4.78%
		April 1	10.0559	27%	0%	32.95%	83.35%	8.83%

All definitions display only modest reduction in error and low median RPSS scores for a March 15 issue date, but performance in both metrics increases consistently with later issue dates. Median RPSS for the five earliest and five latest onset dates generally—but not always—increase with later issue dates, suggesting that although later issue dates result in more accurate forecasts, uncertainty bounds are not substantially reduced. This is particularly true for the threshold method, in which RPSS scores for late years actually drop for a May 1 issue date (Table 2-3) and for which error bars do not noticeably narrow with later issue dates (Figure 2-5). Later issue dates also correspond to occasional missed onsets, in which the forecast is issued after onset occurred. This is also a notable problem for the threshold definition, in which 5% of historical onsets occurred before April 15 and 23% occurred before May 1. In contrast, no onsets were observed before April 1, for any definition, over the entire study period.

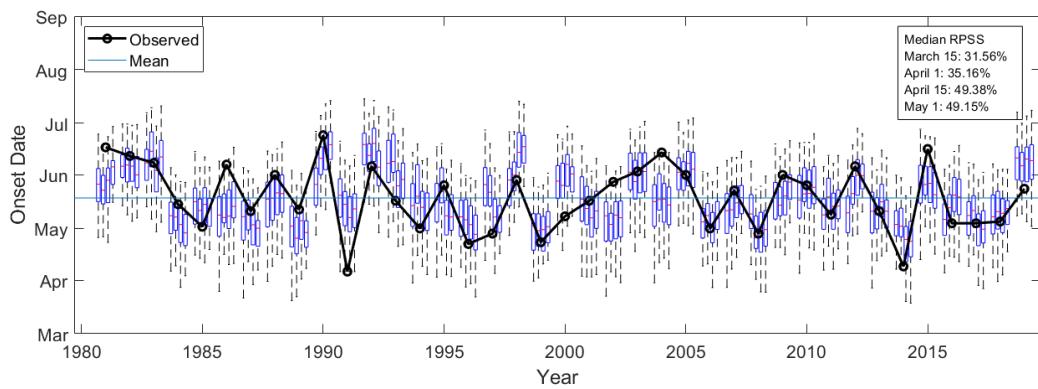


Figure 2-5. Ensemble hindcasts, threshold definition, for March 15 (leftmost bar) to May 1 (rightmost bar) issue dates, along with observed onset date (black line and circles) and climatological mean onset date (blue line)

In contrast to the threshold definition, the yearly definition demonstrates greater skill in nearly all metrics, while reducing the risk of issuing a forecast after onset has occurred (Table 2-3, Figure 2-6). Confidence intervals also tend to narrow, albeit slightly, with later forecast dates, which is less apparent for the threshold definition.

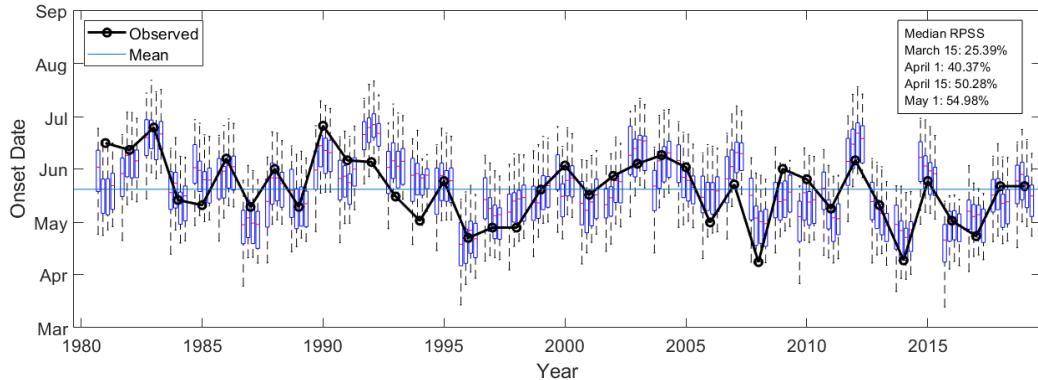


Figure 2-6. Ensemble hindcasts, yearly definition, for March 15 (leftmost bar) to May 1 (rightmost bar) issue dates, along with observed onset date (black line and circles) and climatological mean onset date (blue line)

The window definition demonstrates the least skill among definitions, and confidence intervals only marginally narrow with a later issue date (Figure 2-7). Predictions from later issue dates (April 15 and May 1) are also conditioned on the same set of pre-April predictors, as potential additional predictors from April were eliminated by GCV, thus resulting in no change to the forecast.

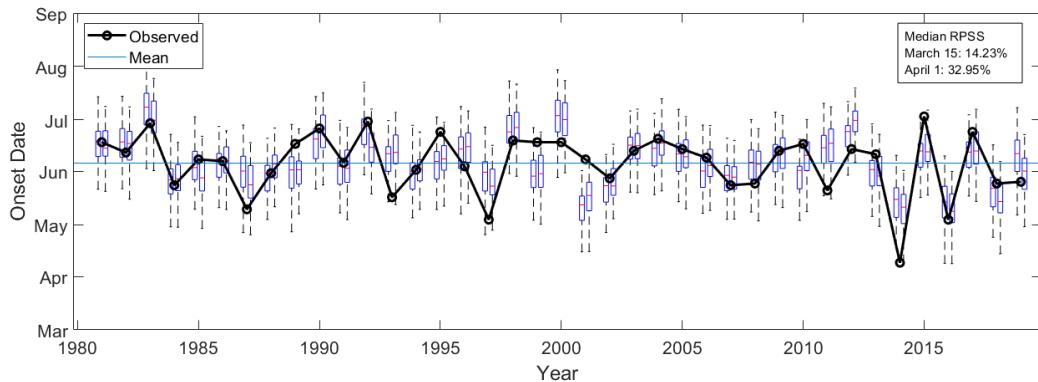


Figure 2-7. Ensemble hindcast, window definition, for March 15 (left bar) and April 1 (right bar) issue dates, along with observed onset date (black line and circles) and climatological mean onset date (blue line)

2.3.3 Qualitative forecast results

Qualitative forecasts for the threshold definition show general improvement with later issue dates, particularly by reducing the rate of extreme misses (Figure 2-8). Correct classification also tends

to increase increases with later issue dates for the PLS regression (from 35% to 54%) between March 15 and May 1, but the improvement is marginal for the random forest method. Overall, the PLS method seems to outperform the random forest method in both correct classification and avoiding extreme misses.

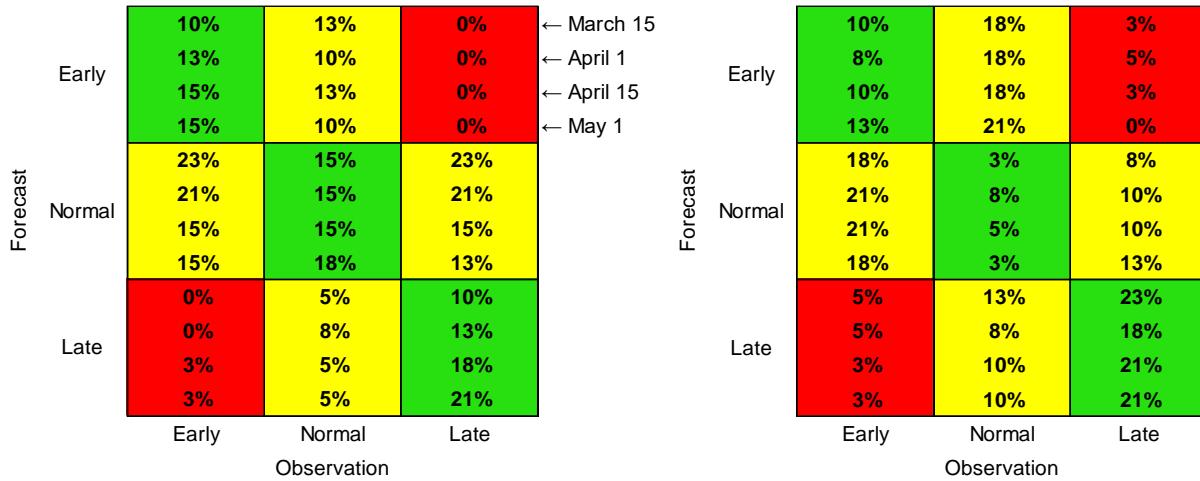


Figure 2-8. Qualitative hindcasts, threshold definition, using PLS regression (left) and random forest classification (right) for March 15 (topmost value for each square) to May 1 (bottommost value) issue dates. Values indicate percentage of years for which hindcasts fell in each category; a perfect hindcast would contain 33% for each value in the diagonal (green) boxes, and 0% in all other boxes

Relative to the threshold definition, qualitative forecasts for the yearly definition have more correct classifications (54% to 70% for the PLS regression and 31% to 62% for the random forest) as well as fewer extreme misses (Figure 2-9). This complements the results of the quantitative forecasts, for which the yearly method is also generally superior.

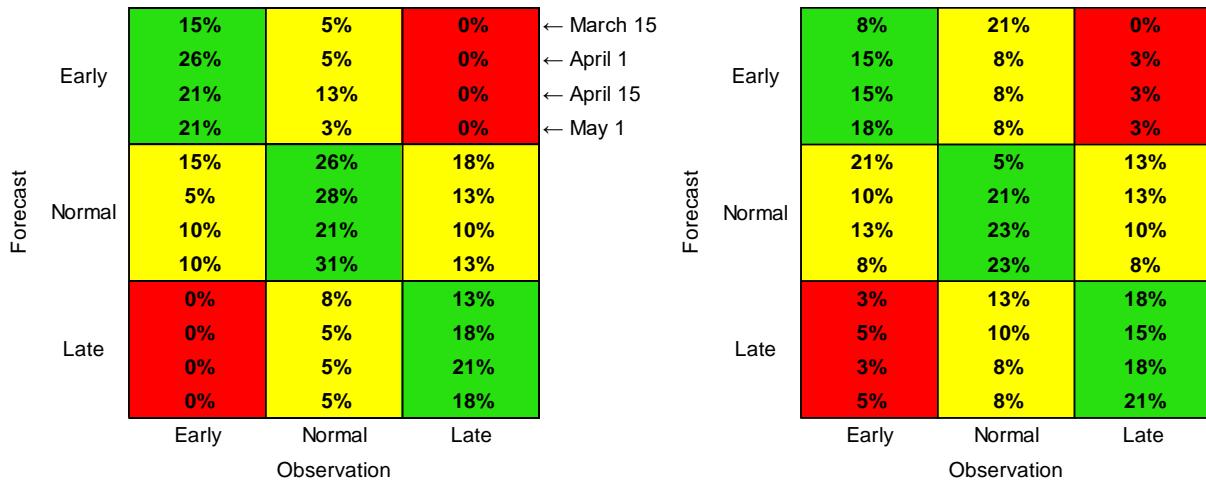


Figure 2-9. Qualitative hindcasts, yearly definition, using PLS regression (left) and random forest classification (right) for March 15 (topmost value for each square) to May 1 (bottommost value) issue dates

Finally, the window method performs moderately well in classifying categories for both the PLS and random forest methods, although both result in several extreme misses (Figure 2-10). Notably, the random forest method slightly outperforms the PLS method in correct classification rates (60% versus 56%, respectively for March 15, and 60% vs 59%, respectively for April 1).

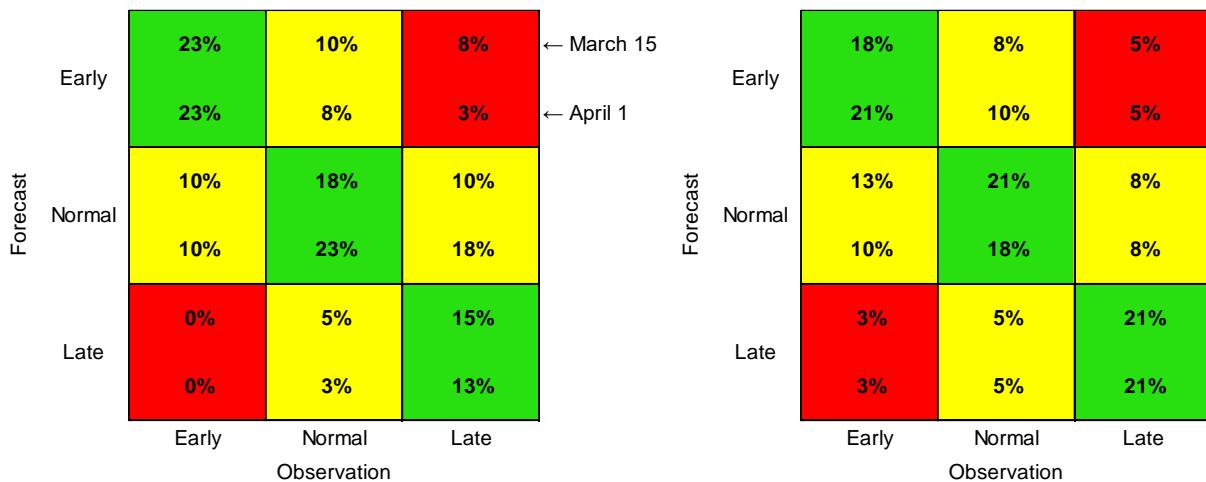


Figure 2-10. Qualitative hindcasts, window definition, using PLS regression (left) and random forest classification (right) for March 15 (top value for each square) and April 1 (bottom value) issue dates

Performance between post-PLS regression classification and direct classification via random forests was mixed; classifying post-PLS slightly outperformed the random forest model for the threshold and yearly definitions and reduced the rate of extreme misses for the window definition, however the random forest model was superior in correct classification for the window definition. In all three definitions, however, random forest classification results in more extreme misses. Thus, there may be scope for considering both models simultaneously.

2.3.4 Application to planting dates

Planting dates vary widely from year to year, with Criterion A usually the earliest and Criterion C the latest (Figure 2-11). The spread agrees well with the surveys, which centered 2018 planting dates around late May with several weeks to either side and centered 2019 planting dates around early June.

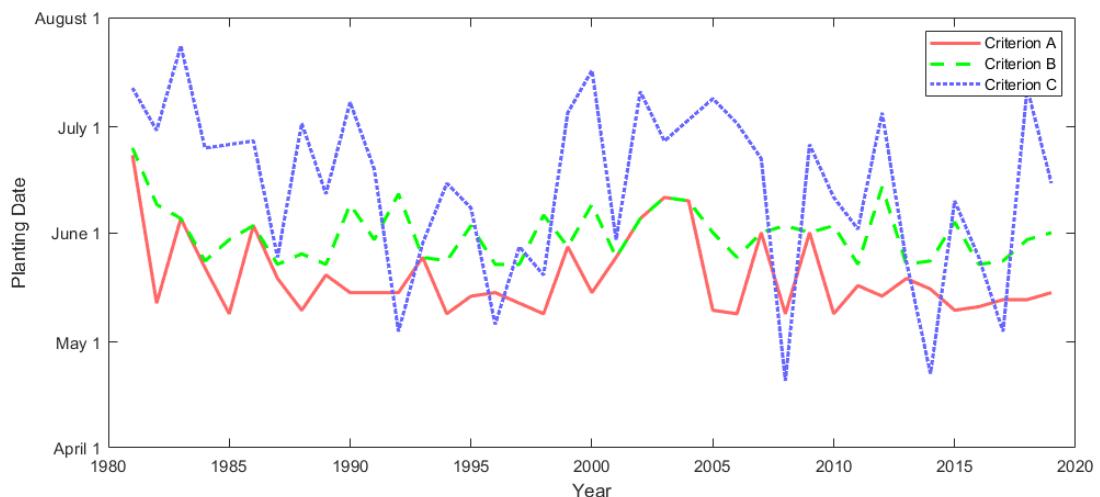


Figure 2-11. Historical planting dates for each criterion

For early-season planting, a primary hazard to farmers is a false onset, in which farmers plant during a wet spell that is followed by a dry spell (Kipkorir et al., 2007; Tadross et al., 2009). Indeed, farmers in the Koga area have specifically requested onset forecasts in order to properly

prepare for planting. To test the utility of the forecasts to this end, a planting date occurring greater than 7 days prior to onset was defined as a “false start”; planting based directly on the forecast was used to compare to the naïve planting criteria. For example, in 1997, early wet spells triggered planting dates on May 14, June 3 and June 9 for Criteria A, B, and C, respectively, but onsets actually occurred on May 26, May 28, and June 24 for the threshold, yearly, and window methods, respectively. Thus, Criterion A was considered a false start according to all onset definitions, whereas Criteria B and C were only considered false starts for the window definition. For years in which there was a false start given a chosen onset definition and planting criterion, forecasts were found to successfully avoid false onsets by advising a later planting date (Table 2-4).

Table 2-4. Number of historical false onsets based on naïve planting criteria and number of false onsets in same years when planting according to forecasts for various issue dates

		Planting criteria		
		A	B	C
Threshold	Naïve planting	11	2	3
	March 15	5	2	0
	April 1	5	2	0
	April 15	5	2	0
	May 1	4	1	0
Yearly	Naïve planting	14	3	1
	March 15	6	2	0
	April 1	7	2	0
	April 15	9	3	0
	May 1	8	1	0
Window	Naïve planting	27	21	9
	March 15	10	10	5
	April 1	11	10	6

As expected, Criterion A runs the most risk of false onset, while Criterion C experiences the least risk. For all onset definitions and forecast issue dates, at least some false onsets are avoided; later issue dates tend to avoid more false onsets except in the case of the window definition. Because

the window definition tends to specify later onsets in general, it has a very high rate of false onsets, suggesting that it is less useful of an onset definition for maize farmers, although it may be useful for mid-season crops or for other sectors. In contrast, the yearly method substantially reduces the number of false onsets, but this reduction is not monotonic with successive issue dates. The threshold definition, however, results in at least a 50% reduction in false onsets relative to any naïve planting criteria for a May 1 issue date.

2.4 Conclusions and Discussion

This study presents an evaluation of *Kiremt* onset predictions using statistical methods. Three definitions of onset and three lead times are proposed, applying both a PLS regression and random forest classification method. Predictions illustrate moderate skill by improving over climatology, although later prediction issue dates (e.g. April 15 and May 1) result in missing some onsets (occurring prior to the issue date). Ensemble predictions using PLS regression show moderate skill in both average errors and RPSS values, especially for later issue dates. Qualitative predictions are classified using both random forests and PLS regression, however neither was unilaterally superior across all onset definitions. Results compare favourably with dynamic model predictions.

The variable performance of predictions considering onset definition and issue date suggests that onset is perhaps not best defined with a single method or model. Individual sectors – agriculture, pastoralism, reservoir management, etc. – certainly have varying definitions of onset, highlighting the importance of tailored decision-making for forecast-informed management decisions. Likewise, complex, heterogeneous geography or climate can further compound variation between onset definitions, motivating localized modelling approaches.

Although no significant correlation between onset date and *Kiremt* total seasonal precipitation for the Koga region is evident, all definitions display a trend of increasingly early onset (~0.3-0.6 d yr⁻¹). In fact, total precipitation during the core months of June-September has remained stationary, suggesting the *Kiremt* season is getting longer but not necessarily wetter. Correlations between onset and total annual precipitation, however, are significant (Pearson's r of -0.60, -0.26, and -0.51, for the threshold, window, and yearly onset dates, respectively) in the Koga region, although total annual precipitation does not independently demonstrate a statistically significant increasing trend. Thus, precipitation earlier in the year has a greater effect on onset date than does precipitation during the core months of the *Kiremt*.

This study also highlights the importance of including atmospheric variables as seasonal predictors as suggested by Nicholson (2017). Geopotential height is found as an important predictor for all onset definitions; zonal wind is also used to a lesser degree. SST is found to play a limited role in onset prediction relative to other variables; ENSO (i.e., SST in the South Pacific) is ruled out entirely by the GCV method, suggesting that more localized predictors play a larger role in onset variability. Pressure variables near the eastern Mediterranean and Red Sea are featured in all onset models at all lead times, which is consistent with the flood predictors of Jury (2011) and serves to demonstrate the blurring distinction of onset as a climate versus weather event.

In most metrics, forecasts using the yearly definition outperform those using the threshold definition. This may be attributable to its inherent connection with seasonal precipitation, which has been shown to be moderately to well predicted at similar lead times (e.g. Korecha and Barnston, 2007; Alexander et al., 2019), whereas the threshold definition is significantly more sensitive to particular precipitation events, making later issue dates more reliable. Thus sectors that place importance on total seasonal precipitation, e.g. reservoir management, may prefer

anomaly definitions whereas agriculture may favor threshold definitions as they are more closely related to soil moisture (MacLeod, 2018). All onset definitions also tend have higher RPSS scores for extreme events (defined as the earliest five and latest five historical onsets); the yearly definition's reach up to 96% for early and 70% for late. Hence, if abnormally early or late onsets are of particular interest, statistical forecasts, particularly using the yearly definition, may be of particular use.

The trade-off between lead time and performance for the threshold and yearly anomaly definitions also warrants consideration. Including additional climate information and issuing forecasts later (April or May) does not necessitate an improvement in performance. Mean absolute error and median RPSS consistently improve with later issue dates, but RPSS scores for extreme events do not monotonically increase, suggesting an improvement in accuracy but not precision, which may explain the relatively small changes in error bars in Figure 2-5 – Figure 2-7. Planting during a false onset, likewise, is less likely when using forecasts, but rates are only slightly reduced for later forecast issue dates. Missed onsets is also a telling statistic; although a March 15 or April 1 issue date may be inferior in terms of percent reduction in error or RPSS, these issue date always occurred prior to all onset dates (39 years), whereas 3% - 23% of onsets have already occurred by the April 15 and May 1 issue dates. For the Koga case study, an April 1 yearly definition prediction is arguably the best-performing choice, striking a balance between lead time, forecast skill, and no missed onsets, however this is not likely generalizable to other locations.

Regarding utility of these forecasts for agricultural planning, the risk of planting during a false onset is substantially reduced when using later forecast issue dates. Although these later dates are at risk of being issued after onset has occurred, they still are issued before farmers will ever plant based Criteria A and B, and before all but two historical planting dates for Criterion C. Issuing

both earlier (March and April) and later (May) forecasts, therefore, may benefit farmers, as early forecasts can help farmers prepare for the season while later forecasts can inform specifically when to plant. The forecasts thus successfully reduce the risk of planting too early; however, planting too late also presents some risk to farmers by resulting in a shorter the growing season, which is not addressed in this study. This is particularly true for the window definition, which tends to have much later onset dates; this definition may be best suited for other crops that are planted later in the season.

To the authors' knowledge, this study is unique in examining statistical prediction models of *Kiremt* onset in Ethiopia. The results of this study compare similarly with dynamic model onset predictions (see MacLeod, 2018), although there are notable differences. The dynamic model (ECMWF SEAS5) provides a longer prediction lead time (February 1), but is coarser in resolution than CHIRPS (0.25° versus 0.05° , respectively), and exhibits a wide range of bias—from <7 days to >28 days using a window definition with CHIRPS data; the Koga region specifically illustrates a bias of 14-21 days, exceeding the bias stemming from the statistical forecast model (10-13 days)—for the Ethiopian Highlands. Also of note is the dynamic model's ability to forecast cessation with modest skill; a corresponding statistical forecast model was developed by the authors featuring similar and alternate oceanic-atmospheric predictor types, however it produced only very limited predictive skill. While statistical methods can capture nuanced effects of climate signals at the local scale, they may also be inherently sensitive to small changes, evidenced by variations in predictors between issue dates for a given definition. The local focus of statistical models, moreover, limits their application to broad regions; evaluating the PLS model at other locations using the same set of predictors demonstrated limited skill, except at other areas within the western Ethiopian Highlands at similar elevations to Koga. Consideration of both dynamical

and statistical approaches may thus be advantageous, pairing early issue dates with location-specific predictions.

This study also considers qualitative forecasts, which may be uniquely received and acted on by individual stakeholders. Future research should therefore consider effective communication of forecasts, including the use of quantitative versus qualitative information, and extend these methods to agricultural and pastoral decision making to better quantify expected value.

Chapter 3. The role of rainy season onset uncertainty on maize yields in Ethiopia

Adapted from: Lala, J., Yang, M., Wang, G., and Block, P. (2021). Utilizing rainy season onset predictions to enhance maize yields in Ethiopia. Environmental Research Letters, doi:10.1088/1748-9326/abf9c9

3.1 Introduction

The dominance of agriculture in Ethiopia's economy, along with increasing climate variability and change, leaves the country highly susceptible to delays in the rainy season or changes in overall precipitation (Darabant et al., 2020). Indeed, as recently as 2015, delayed onset of rains contributed to widespread drought in Ethiopia, a trend which is most likely to increase over time (Philip et al., 2018). This is particularly relevant for early season crops such as maize, which are planted soon after the rains start to maximize the length of the growing season (Liben et al., 2015). The benefits of early planting, however, can be negated in years of "false onset," in which a wet period that prompts planting is followed by a dry spell that can reduce seedling density or require replanting (Kipkorir et al. 2007; Tadross et al. 2009). Soil moisture in the early stages of crop growth is critical to a good harvest, particularly in dry years (Yang et al., 2021). The intense drought of 2015, for example, was especially acute in that both the timing and intensity of rainfall (i.e., a late start followed by only sporadic rainfall) lead to failed harvests (Philip et al., 2018). Climate change is likely to increase the frequency of such dry seasons (Philip et al., 2018), prompting a need for both quality forecasts and adaptive management practices. Farmers have already begun to adapt to climate change through shifting of cropping seasons, adoption of new cultivars, and more flexible

timing of planting dates to align with precipitation (Darabant et al., 2020), but the need for clear, useful forecasts remains.

In Chapter 2, we detail how a statistical forecast of rainy season onset timing performs comparably to dynamic models (e.g., MacLeod, 2018) and significantly reduces error relative to a naïve climatology forecast. Despite this advancement, minimal investigation into the impacts of false onsets has been conducted, prompting further study. Yang et al. (2020) adapted the Decision Support System for Agrotechnology Transfer model (DSSAT; Hoogenboom et al., 2012) to Ethiopia and find that precipitation is significantly correlated with maize germination, motivating evaluation of rainy season onset forecast impacts. The forecasts associated with Chapter 2 (Lala et al., 2020) and MacLeod (2018) have shown promise in predicting false onsets; however, their utility has not been quantified in terms of potential yield gains. This chapter quantifies yield gaps associated with suboptimal maize planting times by combining rainy season onset predictions with a dynamic crop model to highlight the value of pre-season information for agricultural decision-making.

Ethiopia's heterogeneous geography and climate contributes to a wide range of cropping seasons and farm practices, with eight distinct *agro-ecological zones* (AEZs) delineated by elevation, hydroclimatic regimes, and moisture zones (FAO, 2010; Figure 3-1). This study considers yield gaps from maize planting only in the summer *Kiremt* rainy season, during which ~90% of maize production in Ethiopia occurs (Taffesse et al., 2012). The spring *Belg* rains, although occurring in much of the country, are generally not drivers of maize production, except in the far south, and are not included in this study. Even within the *Kiremt*-dominated zones, however, precipitation regimes can range from dry and variable to wet and reliable, suggesting that forecast value may

vary depending on location. This research aims to quantify this value, allowing for the targeted use of forecasts to increase user benefits (Alexander et al., 2019).

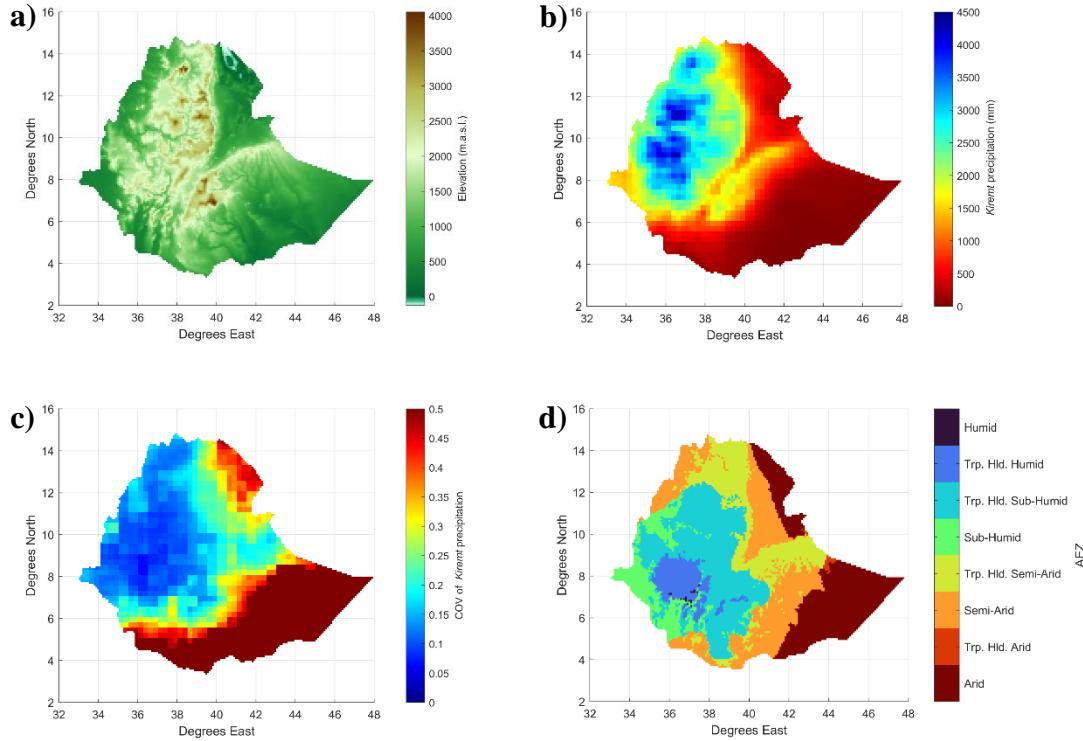


Figure 3-1. a) Elevation (m.a.s.l.), b) mean *Kiremt* (June-September) precipitation (mm), c) coefficient of variation of annual *Kiremt* (June-September) precipitation, and d) AEZs of Ethiopia

Farmers typically aim to time maize planting with rainy season onset, but clear demarcation of the onset can be difficult. Definitions often include a precipitation threshold (e.g., 25 mm of rain over 3 days, and no dry spells of 8 days thereafter; Segele and Lamb, 2005) to represent optimal soil moisture for planting. This methodology, however, may be hampered by limited long-term station data and remotely sensed gridded precipitation datasets not well suited for daily-scale thresholds, particularly in representing of dry days (MacLeod, 2018). As an alternative approach, the cumulative rainfall anomaly relative to some long-term mean is calculated and onset defined as the minimum of this anomaly over a given year (see Liebmann and Marengo, 2001; Dunning et al., 2016; MacLeod, 2018). These methods also have the advantage of avoiding “false onsets” by

considering dry spells throughout the year or season of interest (MacLeod, 2018), and they often differ little from threshold-based methods in practice (Dunning et al., 2016; Lala et al., 2020). Furthermore, although planting at onset is assumed to be optimal for maximizing yields by ensuring a long growing season, the true onset date is unknown prior to the season. Onset predictions, however, have demonstrated reasonable skill in predicting onset, with mean errors ranging from 9-12 days based on localized empirical studies (Chapter 2; Lala et al., 2020) or a wider 0-28 days using regional dynamic model studies (MacLeod, 2018). Defining onset in a way that is useful for adaptive planting approaches, as well as for prediction, therefore constitutes another goal of this research. Finally, although forecast-informed outcomes may represent improvements over the baseline scenario, there is long-standing evidence that farmers, as a whole, are risk averse (Lins et al., 1981; Bar-Shira et al., 1997) and may be reluctant to adopt forecast information. Given this, agricultural extension efforts may be hampered if interventions include losses as well as gains (Yesuf and Bluffstone, 2007). Minimal year-to-year variance in production, therefore, is also assumed valuable to farmers. Accordingly, this study incorporates variance and risk aversion to more accurately represent farmer preferences.

3.2 Methodology

3.2.1 Defining onset and planting dates

Remotely sensed gridded precipitation datasets, which provide for larger, country-wide studies, are best paired with anomaly-based onsets (MacLeod, 2018); however, the exact period for which to calculate anomalies is not necessarily clear. This study therefore investigates two variants of an anomaly-based onset: (1) a *yearly* method, in which the cumulative anomaly is calculated for each year relative to the long-term daily average (Dunning et al., 2016; Supplementary Figure A1), and (2) a *window* method, in which anomalies and daily averages are computed over a window of

interest (in this case, June–December to avoid the earlier *Belg* rains that occur in some parts of Ethiopia; MacLeod, 2018). Precipitation data is taken from the Multi-Source Weighted-Ensemble Precipitation dataset (MSWEP; Beck et al., 2017a) from 1979–2014 at three-hourly temporal and 0.25° spatial resolutions. MSWEP has been shown to accurately capture precipitation relative to other gauge-corrected precipitation datasets globally (Beck et al., 2017b; Beck et al., 2019).

Although the exact planting date clearly relies on a variety of factors, the UN Food and Agriculture Organization maintains data on the typical planting month for many crops, including maize, for the eight AEZs across Ethiopia (FAO, 2010). We assume, as a baseline, that farmers will plant on the first day of the typical planting month in which soil moisture and temperature are optimal; planting is triggered when surface soil moisture is above 40% of saturation and temperature is above 10°C (Yang et al., 2020). As the temperature threshold is nearly always reached by the FAO planting month, the effective trigger is soil moisture alone. The *window* definition of onset is not used to inform planting dates directly; rather, it is used to delineate areas for which the *Kiremt* season is the dominant season for maize planting. In areas where the median *window* onset precedes 1 August, the *Kiremt* season dominates; *Belg*-dominated areas in the southwestern part of the country are not evaluated in this study (Supplementary Figure A2). Mean onset dates from the *yearly* definition approximately match that of the FAO planting months (Supplementary Figure A3) and that of the threshold-based definition of Segele and Lamb (2005), thus, it is considered here as the preferred onset-based planting date. As a further refinement, the onset date is limited to no more than one month before or after the typical calendar month for planting (e.g., for a default planting month of May, onset-based planting is limited to 1 April – 31 June). This ensures reasonable planting dates even in extreme years.

To investigate the value of forecasts for increasing yield, this study compares a realistic *forecast-informed* planting date with the true onset date. For each grid cell (g) and year (i), we generate synthetic forecasts ($j=1-30$) by randomly sampling an error (ε) from a normal distribution with a mean of 0 and standard deviation of 12.5 days, corresponding to a mean absolute error of 10 days (i.e., comparable to average errors from MacLeod, 2018 and Lala et al., 2020) and add it to the true onset date (Equation 1). If a forecast planting date falls before the baseline date, yet soil conditions are too dry (defined in the following section based on estimates from the crop model), the baseline date is assumed for planting. The forecast-informed planting (*FIP*) date is thus defined as:

$$FIP_{g,i,j} = \max \left\{ \begin{array}{l} baseline_{g,i} \\ onset_{g,i} + \varepsilon_{g,i,j} \end{array} \right. \quad (3-1)$$

To test adaptive strategies, this study employs three different sets of planting criteria over the study period of 1979-2014:

1. A baseline criterion representing historical yields, based on the first instance of optimal soil moisture within a given planting month as set by FAO (FAO, 2010), and corresponding to the criterion used in the calibrated DSSAT crop model (Yang et al., 2020).
2. Planting during the true onset date (the *onset-informed* approach), defined using the *yearly* method (Dunning et al., 2016) and limited to one month before or after the FAO planting month.
3. A *forecast-informed* planting date, in which the true onset date is perturbed by introducing random error to create a prediction comparable with the onset forecast accuracy of MacLeod (2018) and Lala et al. (2020). The forecast is disregarded if it occurs before the baseline planting date, since farmers will not plant in dry soil. Thirty sets of simulations,

each with different sets of forecasts, are evaluated, and resulting yields are averaged for each location.

3.2.2 Modeling maize production

To test the effects of different planting criteria, we use a DSSAT model calibrated to observed maize yields based on Ethiopia's Agricultural Sample Surveys from 2004-2013 (Yang et al., 2020). The model simulates growth and development for a variety of crops, including site-specific data such as particular cultivars and levels of fertilizer. Calculations are run on a 5 arc-minute distributed grid system, taking precipitation data from MSWEP, temperature and solar radiation data from ERA-Interim (Dee et al., 2011), and soil data from the Global High-Resolution Soil Profile Database for Crop Modeling Applications, version 2.5 (IRI et al., 2015). Calibration and validation are conducted at the AEZ scale; genetic coefficients are identical across each cell within a given AEZ, with a goal of minimizing both absolute bias (simulated minus observed yield) and relative bias (absolute bias divided by observed yield). After calibration, Yang et al. (2020) found the model to demonstrate generally low bias for maize yields across Ethiopia.

Yields from individual model simulations based on the three planting criteria (baseline, onset-informed, and forecast-informed) are averaged by *woreda* (district-level administrative divisions) for comparison. To explore different perspectives on improvement over the baseline scenario, we employ common risk-informed metrics from economics (expected utility; von Neumann and Morgenstern, 1944) and finance (average value at risk, a.k.a. expected shortfall; Artzner et al., 1998), such that three evaluation metrics are used:

1. Mean yield: $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, where y_i is the yield in the i th year and n is the number of years in the study period (1979-2014).

2. Expected utility: $EU = \frac{1}{n} \sum_{i=1}^n \sqrt{y_i}$, demonstrating decreasing absolute risk aversion, which is empirically verified for most farmers in Ethiopia and worldwide.
3. Average value at risk: the average value of the bottom $\alpha\%$ of annual yields, where α is set to 10% in this study.

These metrics are compared across planting scenarios to infer the relative benefit of onset-informed and forecast-informed planting to the baseline planting both spatially and temporally.

3.3 Results and Discussion

Onset-informed and forecast-informed planting generally return higher yields in the northern, western, and northeastern parts of the country, in comparison to the baseline scenario, while the central highlands show little or no improvement (Figure 3-2, Figure 3-4, and Figure 3-5). This broadly reflects the availability of precipitation in each region: those that receive more sporadic rainfall (e.g., the Tigray region) tend to do demonstrate large improvements, while the central highlands, which receive substantial precipitation, do substantially worse due to a mismatch between climatological onset and optimal soil moisture. Overall, the share of area that demonstrates improvement over baseline ranges from 73% to 81%, depending on the planting strategy or performance metric (Table 3-1). The baseline scenario more closely matches observed yields in most areas than do the informed approaches (Figure 3-3), suggesting it more properly represents current planting practices. Across scenarios, bias is most pronounced in the east (semi-arid AEZ, generally overestimating yields), whereas much of the rest of the country exhibits low bias, especially for the baseline approach. The informed approaches exhibit lower yields relative to both the baseline and observed data in the central highlands (tropical highland humid and tropical highland sub-humid AEZs; Figure 3-4), suggesting that onset-based planting strategies may not necessarily be ideal or representative of regions with plentiful moisture. On a year-to-year

basis, the forecast-informed and onset-informed planting scenarios closely match, irrespective of performance, compared to the baseline (Figure 3-5), suggesting that forecast errors are minimal and/or do not translate into notable differences in yields for most years. Regardless of overall gains or losses in mean yield, the onset- and forecast-informed approaches reduce interannual variance in yields, eliminating the drastically poor outcomes due to a delayed onset. This is particularly notable in the case of the tropical highland example (Figure 3-5b), in which the adaptive approaches averted crop failure due to the 1999-2000 and 2008 droughts despite having lower yields in most years relative to the baseline.

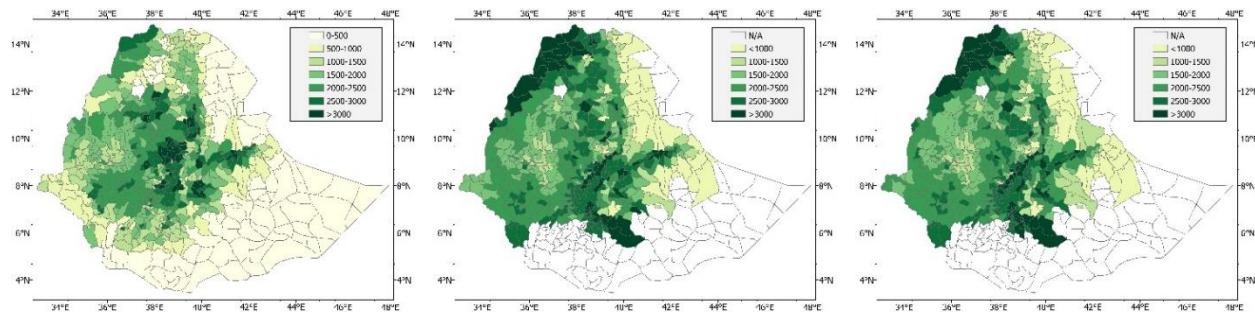
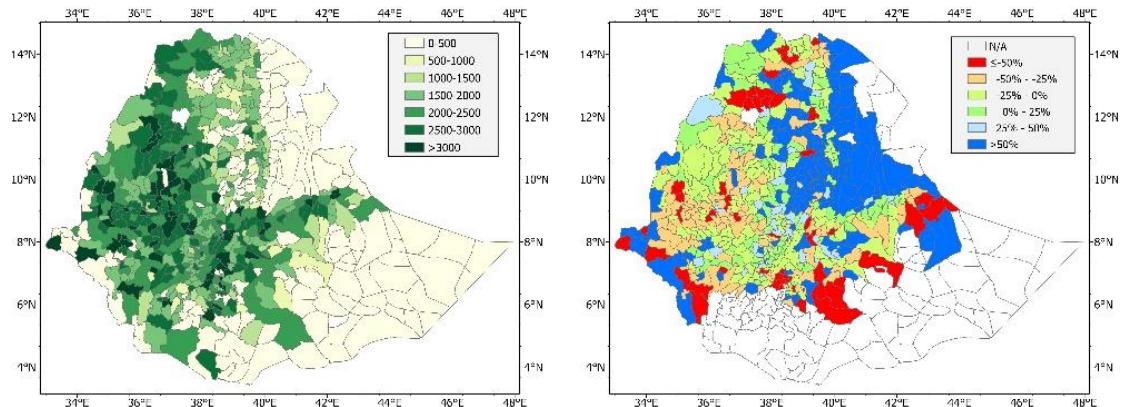


Figure 3-2. Mean yields (kg ha^{-1}) by woreda, 1979-2014, for the baseline (left), onset-informed planting (center), and forecast-informed planting (average of all simulations; right) scenarios



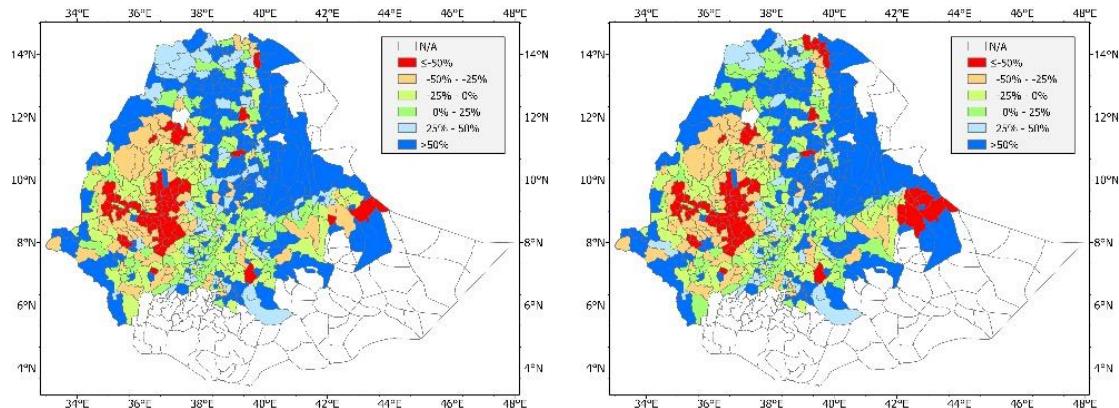


Figure 3-3. Mean observed yield, 2004-2013 (upper left), and relative biases of mean yields relative to observed yields, %, for the baseline (upper right), onset-informed (lower left), and forecast-informed (lower right) scenarios

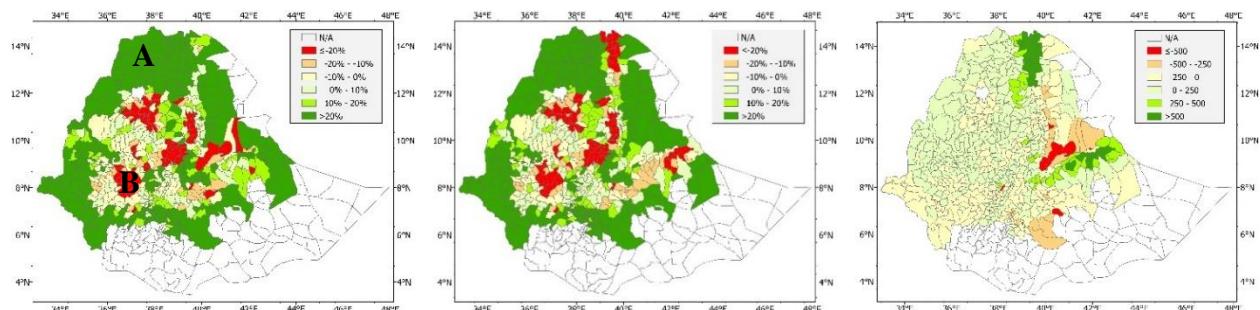


Figure 3-4. Improvement in mean yields for onset-informed planting over baseline (%), left), forecast-informed planting over baseline (%), center), and onset-informed over forecast-informed planting (kg ha^{-1} , right)

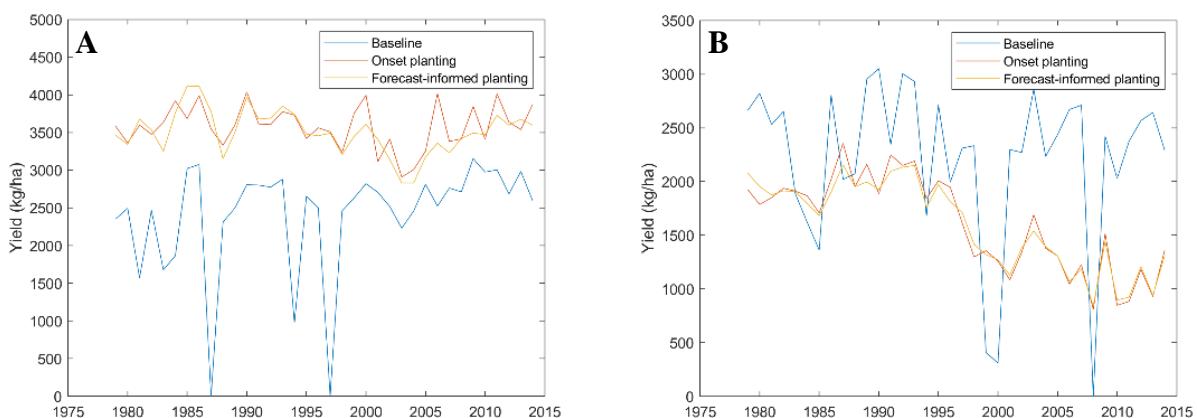


Figure 3-5. Example yields, 1979-2014, from two woredas marked in Figure 3-4: (A) semi-arid climate, and (B) tropical highland humid climate, showing spatial heterogeneity in onset planting benefits

Risk-informed metrics, such as expected utility and average value at risk, demonstrate similar spatial patterns to that of mean yield, with improvements in the north, west, and northeastern parts of the country (Figure 3-6 and Figure 3-7). However, due to the lower overall interannual yield variance of the forecast-informed and onset-informed approaches relative to the baseline, decreases in performance in the central highlands are less stark. In terms of benefits to farmers, therefore, the relative advantages of interannual reduction in yield variance may offset mean yield losses from adaptive approaches, since farmers are likely to weigh severe losses (present in the baseline approach) more highly than overall gains. Regarding average value at risk, α —the bottom percentile of yields from which to average (this study considers values from 5%-30%; Figure 3-7 presents results for $\alpha = 10\%$)—plays a small role; lower levels of α tend to demonstrate slightly better performance (i.e., yields in the worst years improve the most), but spatial patterns are similar regardless of α level.

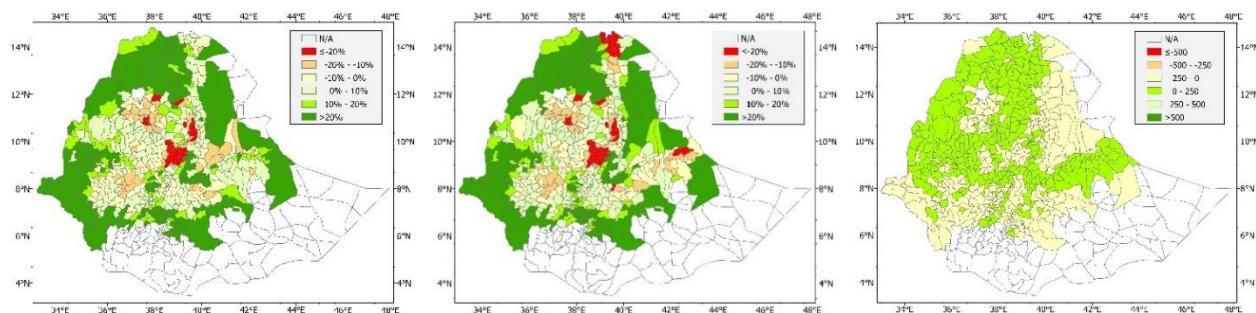


Figure 3-6. Improvement in expected utility for onset-informed planting over baseline (%, left), forecast-informed planting over baseline (%, center), and onset-informed over forecast-informed planting (kg ha^{-1} equivalent, right)

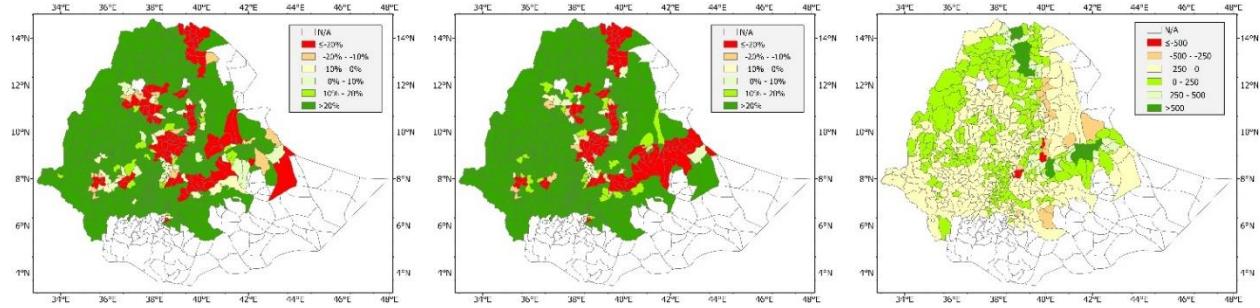


Figure 3-7. Improvement in average value at risk, $\alpha = 10\%$, for onset-informed planting over baseline (%), left), forecast-informed planting over baseline (%), center), and onset-informed over forecast-informed planting (kg ha^{-1} , right)

Table 3-1. Share of area (%) that demonstrates improvement over baseline by planting strategy and performance metric

	Mean yield	Expected utility	Average value at risk ($\alpha=10\%$)
Onset-informed	75%	79%	79%
Forecast-informed	73%	75%	81%

In general, onset-informed and forecast-informed planting results in large yield gains in drier areas, including most of the semi-arid, tropical highland semi-arid, and sub-humid AEZs. The central highlands (tropical highland humid and tropical highland sub-humid AEZs), which tend to have more substantial and reliable rainfall, demonstrate minimal gain—or even loss—when adopting these alternative planting approaches. This may be attributable to the nature of the *yearly* onset definition; for anomaly calculations based on long-term mean precipitation, wet areas by definition must receive more absolute rainfall to trigger onset. Indeed, onset in these wet areas tends to be several weeks *later* than the typical planting date (Supplementary Figure A3). Moreover, this difference appears to be increasing due to climate change, which may explain the slight downward trend in yields (Figure 3-5); evidence suggests that farmers have already begun to alter planting times to counter this pattern (Darabant et al., 2020). Mirroring this, in the eastern parts of the

tropical highland semi-arid AEZ, a drier area where onset is generally several weeks *earlier* than the typical planting time, the alternate planting approaches also result in lower yields. For most other regions, onset corresponds well to typical planting dates and yield differences are not as stark. Such limitations in the application of anomaly-based onset definition (as opposed to an absolute threshold of precipitation) must therefore be considered, despite their advantages for large-scale studies using gridded precipitation datasets (MacLeod, 2018). Further, in moisture-reliable areas, other factors, such as the expected cessation of the rainy season and amount of solar radiation, may play a larger role in determining both yields and farmers' planting decisions (Yang et al., 2020; Alexander and Block, 2021). We note, however, that onset- and forecast-informed planting may still offer benefit by reducing inter-annual variability in yields, as evidenced by the slightly less severe losses as measured by expected utility and average value at risk.

In most parts of Ethiopia, the onset- and forecast-informed approaches are comparable, suggesting that modest forecast errors are outweighed by other factors (e.g. solar radiation and temperature; Yang et al., 2020). In the eastern parts of the country, however, differences between the two informed approaches tend to be more stark. Mean yields and average value at risk are somewhat higher for the onset-informed approach in the tropical highland semi-arid AEZ, while they are lower in the semi-arid (non-highland) AEZ. These same regions tend to have more extreme precipitation variance across years, along with generally low yields overall, which may indicate a lack of precision in the model rather than true differences in the onset- vs forecast-informed approaches. Indeed, absolute differences in yields across scenarios—rather than relative differences—are rather small in the eastern parts of the country (see Appendix of Lala et al., 2021a).

Despite the main focus of this study being on changes in mean yield and interannual variance, it should also be noted that inter-simulation variance provides some insight into agricultural risk. Despite the artificial nature of the forecasts (i.e., they draw from the same error distribution regardless of location), there is still considerable variance between simulations depending on the year. Overall, averaging across all Ethiopia, the highest levels of variance in yields across forecast simulations tend to occur in 1992, 1994, and 2002-2003, which correspond to widespread drought years. This suggests that, in drought years, slight differences in planting date have a stronger effect; future studies may further elaborate on this aspect of risk by considering location-specific forecasts that avoid the spatial homogeneity in forecast error assumed in this study.

Although the large improvements seen in the dry and drought-prone north are partially due to consistent yield gains relative to the baseline across years, some gains are also a result of the more flexible nature of the onset- and forecast-informed approaches, which allow for planting outside the default planting month. In the baseline scenario, if soil moisture is never sufficient for planting within the set month, the model assumes a “failed season” and no production (e.g., 1987 and 1997 in Figure 3-5a), even if soil moisture may be sufficient in the following months. Although this is a limitation of the model, it also demonstrates that gains from the climate-informed planting approaches are most pronounced in years of unexpected or false onset, which farmers cannot easily anticipate. Nevertheless, very large gains in the far north of the country should be interpreted appropriately.

3.4 Conclusion

This study considers the role of onset-informed and forecast-informed planting strategies in determining maize yields across Ethiopia. Results indicate (1) a large potential for yield gains

using an onset- or forecast-informed planting date, particularly in drier parts of the country. Conclusions in the wet highlands are less clear, showing both benefits and drawbacks of new strategies. On one hand, (2) informed planting dates appear to slightly reduce yields in the wet highlands, suggesting that other factors dominate production. On the other, however, (3) these alternative approaches also tend to reduce variance across years such that risk-informed metrics show a more neutral assessment. Such spatial heterogeneity in yield gaps may be related to the nuances of defining onset; the cumulative precipitation anomaly-based definition used in this study is advantageous for large-scale studies using gridded precipitation data (MacLeod, 2018), but (4) differences in the absolute amount of rainfall required for different climate zones may result in wet areas having onsets that occur well after soil moisture is sufficient for planting, which may in turn lead to an excessively short season by the time of rain cessation. More generally, this study concludes that planting based on reliable rainfall is only part of the optimal solution, and that farmers may improve upon these results by considering forecast information alongside other factors not considered in this study, such as indigenous knowledge (e.g., the appearance of flora and fauna associated with impending rain) or the adoption of new technologies. Indeed, farmers in Ethiopia have already begun to adapt their cropping seasons to a changing climate by considering onset trends (Darabant et al., 2020), making the adaptation efforts described in this study more feasible and possible than ever before.

Chapter 4. Incorporating seasonality into an agro-economic model of Ethiopia

Adapted from: Lala, J., Dereje, M., Zhang, Y., You, L., and Block, P. Incorporating seasonality into an agro-economic model of Ethiopia. In draft.

4.1 Introduction

The agriculture-dominant economy of Ethiopia is driven by two rainy seasons: the *Belg* (February-May), with associated “short rains” which are present in the south and east of the country, and the *Meher* (June-September) main cropping season, associated with the *Kiremt* “long rains” and predominant over almost the entire country and are responsible for 65%-95% of total annual rainfall (Segele et al., 2015; Temam et al., 2019). Although most national production occurs during the *Meher*, most major crops are also produced during the *Belg*, particularly in the south and southeast. In years with near-normal precipitation and climate conditions, productivity is typically stable; however, extreme events can cause significant production shocks, with negative climate shocks outweighing positive gains (Block et al., 2008). Given the potential economic impact of drought, studies on drought risk are widespread; relatively few studies, however, consider economic impacts of droughts at the monthly or seasonal scale (Temam et al., 2019).

The impacts of climate and its connection with economic development are of key concern to researchers and policy makers. Past studies have found a strong correlation between GDP and seasonal precipitation (e.g., Diao et al., 2007; Zhang et al. 2020), while others have found positive correlations between rainy season onset and maize prices (Davenport et al., 2021), suggesting that water shortages impacting crop development can have wide-ranging economic effects. Agro-economic studies of Ethiopia do not simply consider correlations, however; some studies have relied on computational economic modeling to extend scientific findings to real-world policy

impacts. Partial equilibrium models in particular are increasingly useful for testing new interventions to reduce drought risk (Dorosh and Minten, 2020). One study found that a 20% increase in maize yields could reduce domestic prices enough to allow profitable export, drastically improving food security and suggesting that the current export ban on maize could be lifted without negative effects (Yami et al., 2020). Numerous studies have likewise utilized partial equilibrium models to estimate the effects of seasonal forecast adoption on the Ethiopian economy (e.g., Block et al., 2008; Zhang et al., 2020). As with drought risk studies, however, many partial equilibrium models only consider processes at an annual scale, offering no insight into seasonal dynamics, despite the role of seasonal and even subseasonal climate processes on drought and cascading countrywide economic effects (Philip et al., 2018; Teshome and Zhang, 2019).

This study extends previous economic modeling efforts to include seasonal climate dynamics utilizing an economy-wide multimarket model (EMM) of Ethiopia, developed by the International Food Policy Research Institute (Diao et al., 2007). Specifically, the modified model incorporates seasonality by disaggregating the modeling process into the two main cropping seasons—the *Belg* and *Meher*—to compare against an annual timestep formulation. In the process, we update the baseline data from c. 2003 to more recently published data from 2014-2017. The result is an up-to-date model to consider risk reduction interventions at appropriate temporal scales, namely the potential effect of *Kiremt* onset forecasts on maize yields (Lala et al., 2021a).

4.2 Methodology

This section describes the Ethiopia economy-wide multimarket model (EMM; Diao et al., 2007)—developed by the International Food Policy Research Institute—and the updates made to the model, including: (1) updating and disaggregating the input data and simulation timestep; (2)

incorporating seasonal climate yield factors (CYFs); and (3) simulating the economic effects of maize yield gains arising from the use of onset forecasts (Lala et al., 2021a).

4.2.1 Baseline model

The Ethiopia economy-wide multimarket model (EMM; Diao et al., 2007) is a partial equilibrium model that simulates 34 agricultural commodities (including livestock) across 56 administrative zones (93 in the updated model) for which data is available. Supply and demand are calculated at the zonal level, with supply being a function of yield and cultivated area and demand calculated as the maximum of a Stone-Geary utility function based on commodity price and zonal per capita income. Price and income elasticities are based on survey data and differ for each zone and commodity. Similarly to a general equilibrium model, per capita income is endogenous and calculated as production revenue divided by population, linking supply and demand. Producer prices are adjusted to represent value added; no intermediate inputs or costs are considered. GDP is calculated as the sum of supply across all zones multiplied by their respective value-added prices; GDP is further disaggregated into agricultural and non-agricultural GDP. The model ignores bilateral trade flows between zones and only considers a single market—Addis Ababa—for trade. Price margins between zonal markets are thus a function of the zone's distance from Addis Ababa. International trade is determined via exogenous global commodity prices, with imports (exports) only occurring when domestic price exceeds (falls below) the import (export) parity cost. Finally, poverty and per capita calorie counts are calculated with data on income levels, consumption, and expenditure. For full details of the model, please refer to Diao et al. (2007).

4.2.2 Seasonal input data and disaggregation

To explore seasonal dynamics, the model timestep and data are disaggregated into two seasons: *Belg* (January-May) and *Meher* (June-December). Note that the dry *Bega* season is distributed into the other two seasons, as minimal agricultural production occurs during this period.

Production data is obtained from Ethiopia's 2014-2017 Agricultural Sample Surveys (Ethiopia CSA, 2017), reported by zone and season. To account for interseasonal storage for a variety of crops, seasonal supply—distinct from seasonal production—is calculated using the survey results of Bachewe et al. (2018). Commodities produced throughout the year—specifically livestock, sugar, coffee, enset, and bananas—are only reported as annual values; for these, we disaggregate into seasonal production according to seasonal consumption rates, as sale and production is generally evenly dispersed throughout the year (Bachewe et al., 2018). Details on farm management practices, including specific inputs—irrigation, fertilizer, improved seed, pesticide, or a combination—is obtained from the Agricultural Sample Surveys (Ethiopia CSA, 2015).

Consumption and expenditure data is from the Ethiopian Household Consumption – Expenditure (HCE) survey (Ethiopia CSA, 2018), reported at the zonal but at an annual scale. To disaggregate into seasons, we follow the survey results of Hirnonen et al. (2015), which reports consumption by month and household type (i.e., urban or rural) for a variety of commodities.

4.2.3 Incorporation seasonal climate variability

Seasonal climate variations have a strong effect on crop production in Ethiopia (Philip et al., 2018; Teshome and Zhang, 2019). Climate variability is incorporated here through the use of climate yield factors (CYFs; Block et al., 2008) for both the *Belg* and *Meher* seasons. CYFs quantify the yield response to water via a process-based crop model and are particularly useful in water-stressed

areas with high spatiotemporal variability such as Ethiopia (Zhang et al., 2020). CYFs are determined for five staple crops: maize, sorghum, wheat, teff, and millet.

CYFs are calculated using gridded monthly climate and geomorphic data. Reference evapotranspiration (ET_0) is calculated using the Penman-Monteith method (Allen et al., 1998) based on elevation, solar radiation, humidity, wind, and temperature data from the University of East Anglia's Climate Research Unit (CRU; Harris et al., 2020). Potential evapotranspiration (ETC) is then determined by multiplying ET_0 by a crop-specific empirical water stress coefficient K_s . Actual evapotranspiration (ETA) is calculated using a soil water balance model (Allen et al., 1998) based on monthly precipitation data from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS, Funk et al., 2014) and soil type maps from the UN Food and Agriculture Organization (FAO-UNESCO, 1988). The CYF is then calculated as:

$$CYF = 1 - K_y \left(1 - \frac{ETA}{ETC} \right) \quad (4-1)$$

where K_y is a crop-specific empirical yield response factor indicating a crop's sensitivity to water stress. The CYF can vary from 0 to 1, with 0 indicating crop failure and 1 indicating no water stress. Other stresses (e.g. fertility stress, pests, etc.) are not considered. The CYFs are computed independently for the *Belg* and *Meher* seasons for each year 1981-2020; for the aggregated annual timestep model, the CYF is calculated from the seasonal-yield weighted average of the *Belg* and *Meher* CYFs. Finally, CYFs are aggregated to the zonal scale by averaging across grid values falling within each zone.

4.2.4 Forecast-informed yield gains for maize

Maize is the most water-sensitive cereal crop in Ethiopia and expresses significant interannual variability. The application of seasonal and subseasonal forecasts for planning, however, has been demonstrated to reduce this variability and potentially increase yields under certain conditions (Zhang et al., 2020; Lala et al., 2021). To explore the economic effects of such interventions, we incorporate the yield gains from Lala et al. (2021) representing the adoption of *Kiremt* onset forecasts for timely planting of maize (Figure 4-1). Given that these yield gains arise from more efficient water use, we incorporate them into the CYF for each year during the *Kiremt* season:

$$CYF_i^{fcst} = CYF_i^{baseline} * \frac{yield_i^{fcst}}{yield_i^{baseline}} \quad (4-2)$$

where i is the year, $fcst$ refers to the forecast-informed strategy, and $baseline$ refers to the baseline approach. These updated CYFs are limited to a maximum value of 1 since they represent only yield gains from reduced water stress. Not all zones demonstrate yield gains or reductions in interannual variability; thus, only zones that typically expect gains in expected utility over the baseline include updated CYFs (Figure 4-1, right).

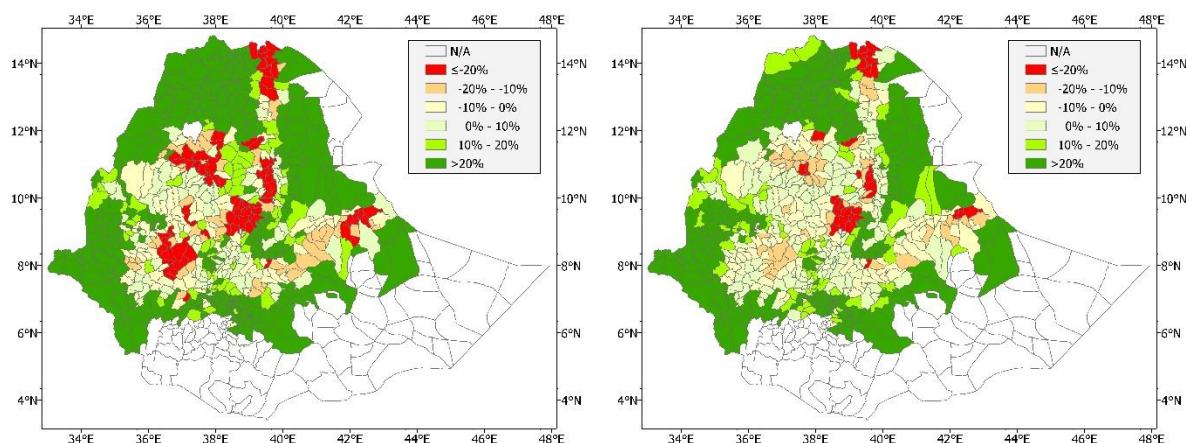


Figure 4-1. Gains in mean yield (left) and expected utility (right) over baseline conditions resulting from forecast adoption (Lala et al., 2021a)

4.3 Results

In general, inter-seasonal climate variability has a greater effect on *Belg* yields than on *Meher* yields, with cultivated area-weighted CYFs for the *Belg* demonstrating a lower mean and higher variance than *Meher* CYFs for all major cereal crops (Figure 4-2). Consistent with Zhang et al. (2020), maize tends be highly susceptible to water stress, exhibiting lower CYFs and greater variance than other crops, particularly in the *Meher* season; contrastingly, teff is drought-resistant with the highest mean CYF and lowest variance in most cases. The considerable variability in *Belg* CYFs highlights the season's particular susceptibility to rainfall failure and its secondary nature in most regions to the longer and more reliable rains of the *Meher* (Rosell, 2011).

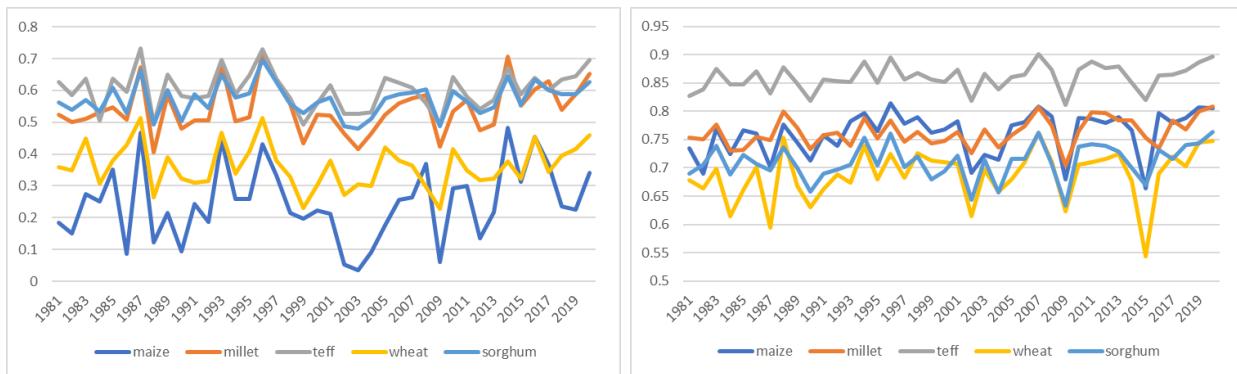


Figure 4-2. Cultivated area-weighted CYFs for *Belg* (left) and *Meher* (right) for five major cereal crops (note the difference in scales)

Commodity prices also vary by season and year, as expected, given the relatively inelastic demand of cereal grains and large interseasonal and interannual variances in yield. Similar to yield (Figure 4-2), price varies the most for maize, particularly for the *Belg* season (Figure 4-3). Price fluctuations are partially stabilized, however, given the availability of storage; on average, approximately 25% of *Meher* cereal production is not consumed or sold until the *Belg*, offsetting some of the extreme variance in *Belg* yields that would otherwise result in drastically higher producer prices (Bachewe et al., 2018). For the forecast-informed approach, prices on average

decrease by 7-10% relative to the baseline, attributable to higher overall yield gains (Figure 4-3); however, the forecast-informed approach also modestly increases interannual variance in yields, owing to a reduction of prices in productive years without a corresponding reduction in unproductive years.

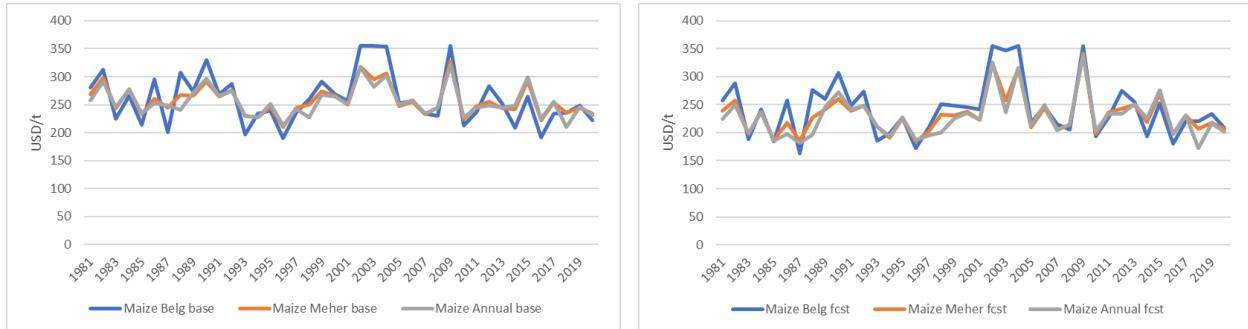


Figure 4-3. National maize producer price for baseline (left) and forecast-informed approach (right)

Consistent with the results in yield and price, GDP and calorie consumption also vary by season, year, and strategy (Figure 4-4). Given the greater production levels and longer duration of the *Meher* (assumed here as seven months vs. five for the *Belg*), GDP is higher than in the *Belg*. Calorie consumption, likewise, is higher in the *Meher*. The forecast-informed strategy demonstrates considerable gains in both measures, increasing both GDP and per capita calorie consumption in nearly all years, particularly in the *Meher* in which the yield gains take place. On average, the forecast-informed strategy increases GDP by 0.9% and per capita calorie consumption by 1.7%, suggesting that this strategy can have broadly positive effects on economic development. Moreover, gains in GDP and calorie consumption are not limited to the *Meher*; rather, the effects of storage and trade ensure that spillover of these gains extend to the *Belg*.

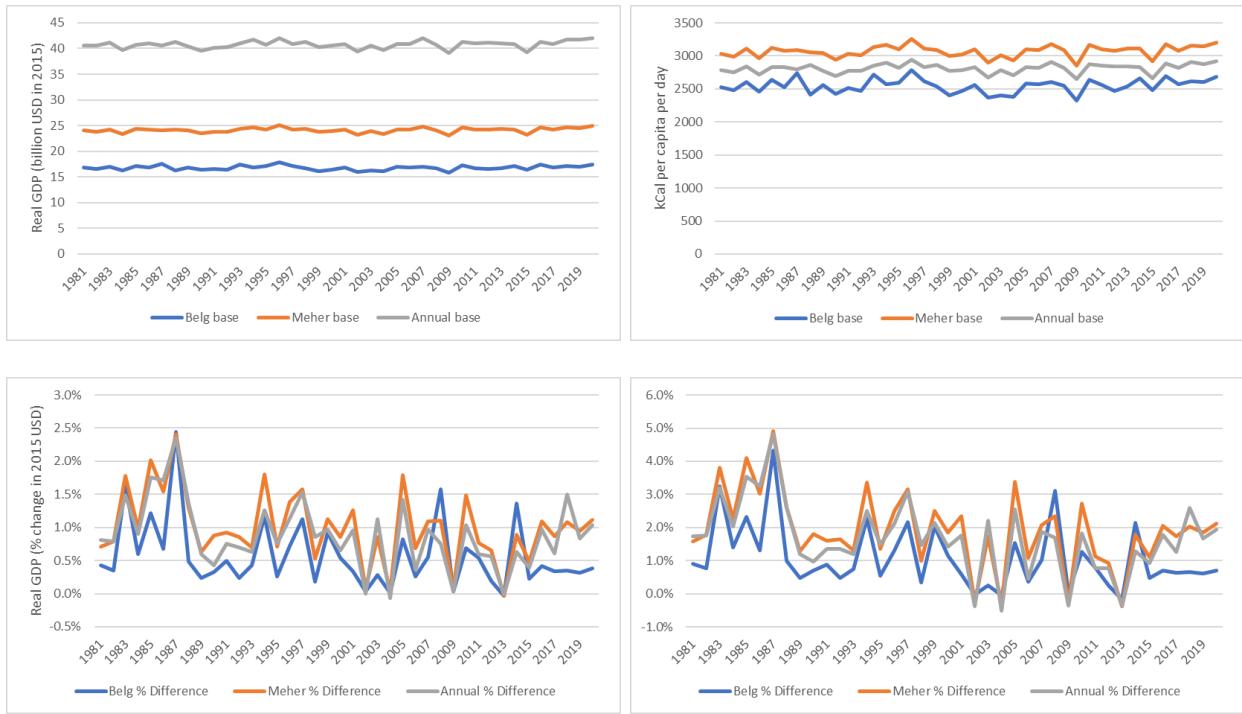


Figure 4-4. GDP (left) and daily calorie consumption per capita (right) for baseline (top), and % gains by using forecast-informed strategy (bottom)

Finally, poverty rates also demonstrate seasonal variance and general improvements using the forecast-informed strategy (Figure 4-5). Again, given the volatility in *Belg* rains and the general predominance of *Meher* production, poverty rates are higher in the *Belg*, especially among rural households, where the poverty rate hovers around 26%. The forecast-informed strategy, on average, reduces both rural and urban poverty, but stronger, consistent reductions are only seen in the urban poverty rate, suggesting that trade dynamics influence the allocation of these gains despite originating in rural areas. This may be attributable to the single market in Addis Ababa, the dominant urban area in which there are no transportation costs for consumption and for which household wealth is less contingent on agricultural production.

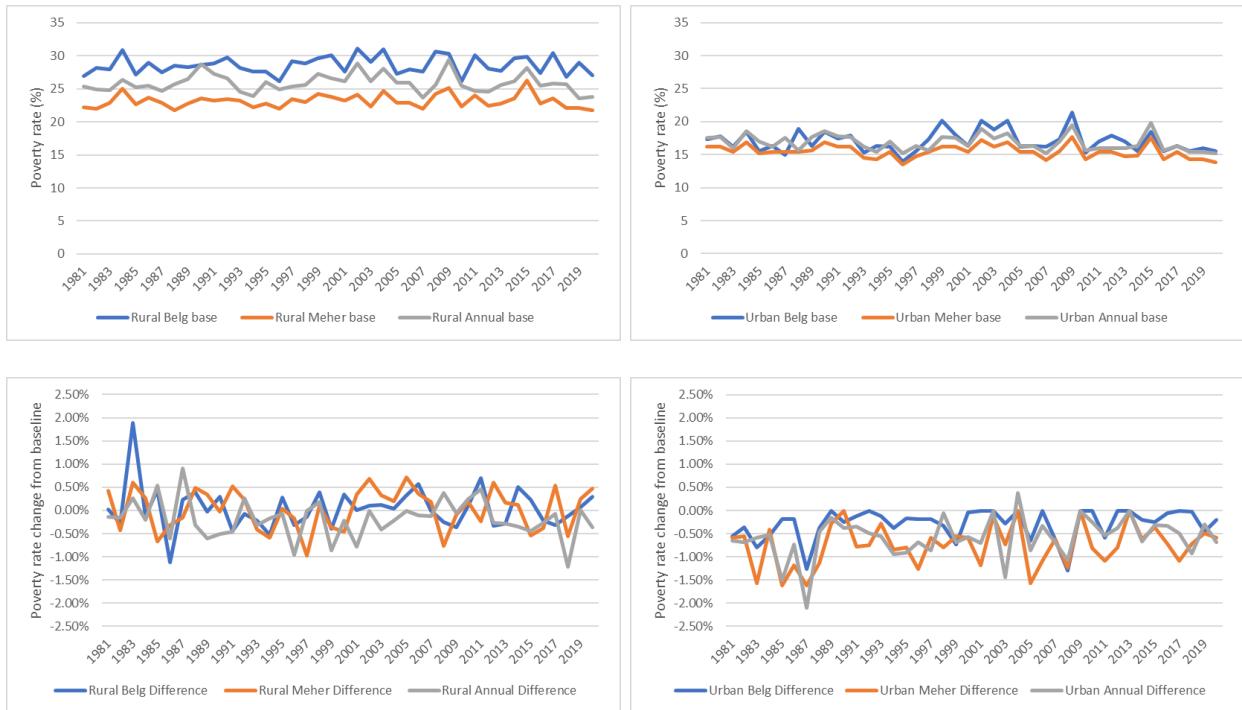


Figure 4-5. Rural (left) and urban (right) poverty rate for baseline (top), and change forecast-informed strategy (bottom)

Along with temporal variance by season and year, gains in GDP per capita also demonstrate spatial variance that roughly matches that of zonal gains in yields (Figure 4-6). In most of the country, GDP gains are small but positive, though they are more substantial in areas for which there are strong yield gains (e.g., Tigray). The Central Highlands, in which the forecast-informed strategy is not adopted due to a lack of yield gains, contrastingly see a *decrease* in GDP per capita, owing to no change in production but an overall decrease in price due to increased supply elsewhere in the country. Seasonally, the *Belg* overall sees the least amount of change in GDP, given that changes in maize supply are only due to extra storage from the *Meher*; the *Meher*, in contrast, demonstrates the starker changes in zonal GDP.

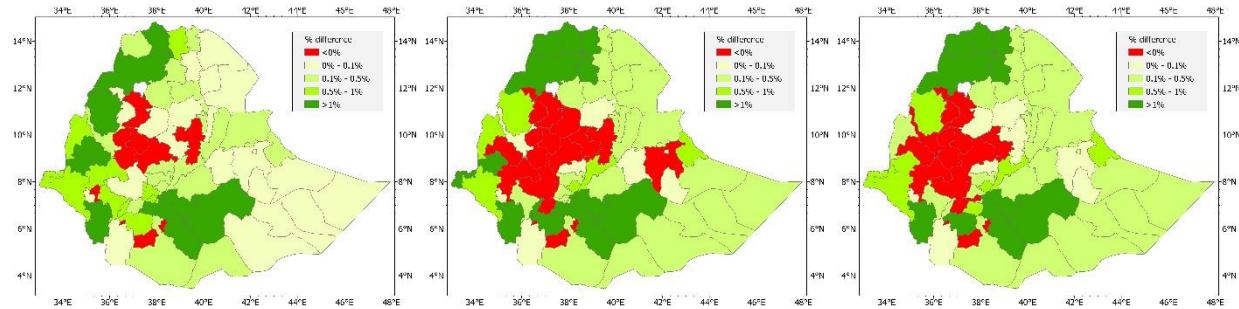


Figure 4-6. Zonal changes in GDP per capita, % difference over baseline, averaged over 1981-2020, for *Belg* (left), *Meher* (center), and annual (right) seasons

4.4 Discussion

By incorporating seasonality into the EMM, we highlight two main findings. First, yield and production vary considerably by season, and resulting economic indicators, such as commodity prices and GDP, mirror that variability. In particular, the *Belg* season tends to express greater variance in yields than the *Meher* season, stemming from the volatility of rainfall during those months, leading to more variance in price. When averaging over all years, *Belg* prices are higher for most crops, again stemming from lower production that is not captured in the annual timestep. Second, the use of onset forecasts to increase maize yield in the *Meher* has reverberating effects across all seasons, leading to overall lower maize prices, higher levels of GDP and per capita calorie consumption, and lower levels of poverty. Such reductions in price are supported by empirical findings on the relationship between onset and maize prices (Davenport et al., 2021), while reductions in poverty via a growth in staple crops is documented in earlier studies (Diao et al., 2007), further strengthening the argument for increased adaptation measures as rainy seasons become more variable in timing. Based on the findings here and in Zhang et al. (2020), we conclude that statistical climate forecasts can play a key role in scenario analysis and risk reduction.

Despite some clear benefits to forecast adoption, however, we caution that the spatial heterogeneity of these benefits causes corresponding heterogeneity in gains or losses in a dynamic system. Specifically, the overall gains in maize yield lead to lower prices, which tend to offset most potential economic gains and actually result in *GDP losses* in the Central Highlands, where maize production remains static. In general, we conclude that the forecast-informed strategy is good for consumers (e.g., lower prices, higher levels of calorie consumption and overall GDP, slight reduction in poverty rates) but has mixed implications for producers (i.e., depending on the magnitude of yield gains, the corresponding decrease in prices may or may not eliminate gains in income). Still, we argue that the overall gains in GDP and calorie consumption, as well as slight decreases in national poverty rate, merit consideration of forecast adoption in a targeted approach.

Given the large number of assumptions inherent in partial equilibrium models, it is worth comparing the results of this study with others using a similar model framework. In general, the EMM produces relatively small price responses to supply shocks, compared to other models. Using a partial equilibrium model of Ethiopia conducted at a monthly timescale, Bakker et al. (2018) found that a cereal supply decrease of 25% yields a 60% increase in producer prices in East Oromia, a major maize growing region vulnerable to drought. In the EMM, by contrast, a decrease in cereal supply of 25% over the mean—which occurred in 1990 in Bale, Oromia—increased producer prices by only 2-25%, depending on the specific cereal grain. Similarly, Yami et al. (2020), considering maize at the annual scale over all of Ethiopia, found that a supply increase of 20% decreases the price by 81%, while a supply decrease of 15% increases the price by 61%. Similar shocks and gains in the EMM—a 13% increase in supply in 2018 and a 15% decrease in 2009—leads to a drop of 18% and increase in 26% in producer prices, respectively. We thus

highlight the sensitivity of these models to different model assumptions, particularly elasticity and consumption habits.

Finally, there are certain limitations and proposed updates to the model that are not addressed in this study. Firstly, the model does not consider the effects of mechanization, which may have a strong effect on the labor force, if not yield and production (Berhane et al., 2017). This is likely to be a major factor in the future of the Ethiopian economy; however, widespread adoption of operational commitments to mechanization are still lacking, with only ~1% of cultivated land mechanized (Ayele, 2021). Secondly, while this study includes irrigation as a technology input, the level of irrigated agriculture within the country is expected to expand significantly above the relatively small levels in the current baseline data. The completion of the Grand Ethiopian Renaissance Dam will bring about a large irrigation district, and rapid adoption of small-scale technologies (e.g., solar and diesel pumps, community reservoirs) nationwide will likely have an effect on crop yields and subsequent economic indicators; past studies looking at long-term irrigation investments could raise yields by up to 40% and modestly decrease poverty rates and food deficits (Diao and Pratt, 2007). Future attention to this trend should thus be considered. Finally, seasonal data collection is often lacking in Ethiopia, preventing this study from considering spatiotemporal effects on storage or seasonal consumption. Households may, for example, store a larger proportion of their crops in good years, yet the proportions stored in the model remain static regardless of year. In both cases, we stress the importance of data availability as necessary for more detailed studies of the Ethiopian economy.

4.5 Conclusion

This study addresses seasonal climate dynamics of the Ethiopian agricultural economy by incorporating seasonality into a partial equilibrium multi-market model. In general, we find that the response to high variability in *Belg* season rains is a corresponding volatility in yield, commodity prices, calorie consumption per capita, and poverty rates; in contrast, the more reliable rains of the *Meher* season result in less volatility. We also test a seasonal forecast-informed planting strategy for maize, finding that the strategy decreases maize prices and poverty rates and increases caloric intake and GDP. Price effects, however, have a partially neutralizing effect on yield gains such that some zones experience no gains—or even losses—in income. Given the assumptions inherent in partial equilibrium models, we compare our results to other models using a similar framework, finding a relatively small effect of supply shocks on prices, suggesting that even small differences in model assumptions can lead to notable differences in results. Overall, future modeling efforts may benefit from similar seasonal analyses of economic systems, so that season-specific interventions can be well-represented, particularly in the rapidly growing field of seasonal and subseasonal forecasting.

Chapter 5. Optimizing forecast-based actions for extreme rainfall events in Peru

Adapted from: Lala, J., Bazo, J., Anand, V., and Block, P. (2021). *Optimizing forecast-based actions for extreme rainfall events*. Climate Risk Management, 34, doi:10.1016/j.crm.2021.100374

5.1 Introduction

The increasing cost of natural disasters—averaging US\$165 billion per year and growing (World Bank, 2014)—is a major challenge to global development. In 2019, floods cost over US\$45 billion and killed nearly 4,500 people; these numbers are expected to increase substantially with climate change (WRI, 2020). Meanwhile, in low- and middle-income countries, disasters can result in poverty traps and total up to 3% of GDP (Borgomeo et al., 2017). Due to advancements in both subseasonal-to-seasonal prediction (WMO, 2018) and financial instruments such as index-based insurance (Skees, 2011), the push for mitigation measures has reached beyond physical infrastructure alone, which may have unintended consequences that actually increase vulnerability (Di Baldassarre et al., 2018). Rather, relief organizations such as the Red Cross have partnered with governments and philanthropic organizations to institute *forecast-based financing* (FbF) initiatives, in which funds for early action are dispersed before a disaster occurs, based on the predictions of a forecast or suite of forecasts. The Red Cross in particular has instituted FbF programs Peru, Togo, Mozambique, Ecuador, the Philippines, and Bangladesh, among others (Coughlan de Perez et al., 2015; Lopez et al., 2017; IFRC, 2020).

Key to the sustainability of FbF programs is the establishment of standardized early action protocols (EAPs), which define the forecasts used and actions taken in advance of a natural disaster. Because this is a relatively new innovation, there are several outstanding issues related to

the use of EAPs, including the use of longer-lead forecasts for prepositioning in addition to short-term forecasts; explicit optimization of the probabilistic thresholds that trigger action to avoid both missed events and false alarms; the incorporation of risk aversion into financial considerations, as opposed to the more traditional cost-loss ratio that assumes a risk-neutral decision maker; and explicit tailoring of global forecasts to local phenomena. This chapter thus aims to answer three main questions: (1) what are the optimal probabilistic trigger thresholds for each lead time and how do they influence each other; (2) what role to economic parameters, such as the cost and benefits of action and the level of risk aversion, play in optimizing the EAP, and (3) to what extent does forecast accuracy affect optimal decisions and outcomes?

5.2 The Peruvian Red Cross early action protocol for extreme rainfall

The Peruvian coast contains more than half of Peru's population but is extremely arid and generally receives only sporadic rainfall, mainly on a seasonal basis (Bazo et al., 2013; Rau et al., 2016). Extreme precipitation events and wide interannual variability in rainfall, however, are not rare and are generally associated with the El Niño – Southern Oscillation (ENSO) phenomenon, with extreme El Niño years bringing potentially catastrophic rainfall and subsequent flooding to the northern coast during the austral summer rainy season. The most recent extreme El Niño, in 2017, affected half a million people in the Piura Region alone; damages were valued at nearly US\$4 billion even though extreme rainfall was mostly confined to a single month (INDECI, 2017). Two decades earlier, the 1997-98 extreme El Niño inundated northwestern Peru with nearly 150 mm of rain over a five-day period at its most intense; over the course of February 1998, over 500 mm of rain fell, and over the three-month period from January-March, nearly 1500 mm fell, killing 366 people and affecting 500,000 more (Peruvian Red Cross, 2018). These extreme events, combined

with the relative predictability of ENSO and support from institutions in Peru, have resulted in the development of early warning systems throughout the country.

Given the presence of both early warning systems and forecast-based index insurance in Peru, the Red Cross Red Crescent Climate Centre and German Red Cross have partnered with the Peruvian Red Cross to institute an EAP (Peruvian Red Cross, 2018) for extreme rainfall in the coastal northwest, as part of a worldwide FbF initiative. The EAP stems from an earlier pilot project, launched in 2015, which was among the first to use the “Ready-Set-Go!” approach of using seasonal, subseasonal, and weekly forecasts in an attempt to bridge the gaps in weather and climate prediction and allow more time for preparatory actions (Bazo et al., 2019). Despite these innovations, there was no evaluation of forecast skill or probabilistic trigger thresholds ahead of time, leading to several false alarms when the project was finally triggered during the 2016 El Niño event. The clear benefits of acting early were thus overshadowed by the costs of acting in vain, prompting the Red Cross to revise its approach and providing some of the research questions addressed in this study.

The current, revised EAP is conditioned on forecasts at seasonal, monthly, and weekly lead times to predict extreme events, with associated actions taken at each lead. Actions fall into three general categories: water, sanitation, and hygiene (WaSH); emergency health; and direct cash transfers. The seasonal and monthly forecasts are used for prepositioning of supplies and coordination of volunteers; specifically, teams for WaSH and health are organized and deployed to at-risk regions, and low-regret supplies such as soap, chlorine tablets, and medicines are purchased. The weekly forecast then serves to activate interventions at a specific community by deploying volunteers, distributing WaSH and health kits, and paying out direct cash transfers to particularly vulnerable

community members. The seasonal and monthly forecasts thus act to provide adequate preparation for a community intervention triggered by the weekly forecast.

The EAP defines extreme rainfall as exceeding the 95th percentile of cumulative rainfall for the seasonal and monthly forecasts, and the 99th percentile of cumulative rainfall for the weekly forecast. Forecasts provide a probability p of exceeding this level, and actions are triggered by setting a trigger threshold probability τ such that $p \geq \tau$. Forecasts are issued one month ahead for the seasonal and monthly forecasts and 2-7 days ahead for the weekly forecasts (e.g. for an issue date of 1 January, the seasonal forecast would predict February-April precipitation, the monthly forecast would predict February precipitation, and the weekly forecast would predict precipitation for 3-7 January; Figure 5-1). Seasonal and monthly forecasts are issued each month and can be triggered independently; weekly forecasts are issued daily and can be triggered only if seasonal or monthly actions have already been triggered for that month.

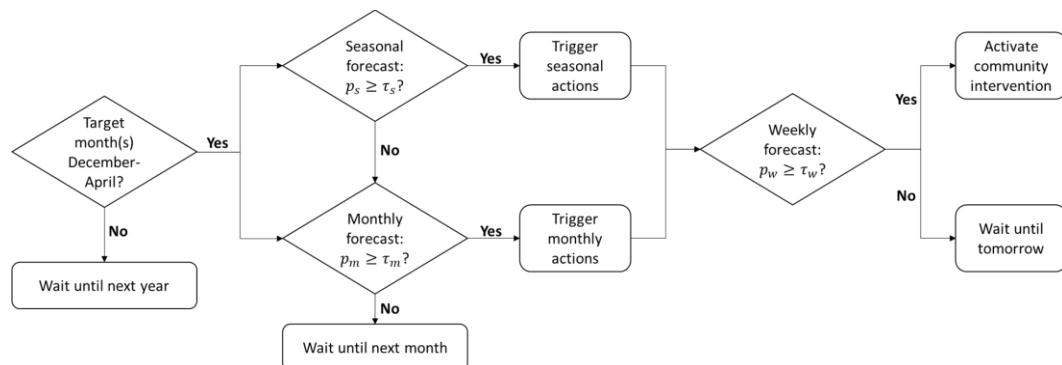


Figure 5-1. Process flow chart of EAP forecast triggers

Currently, forecasts from readily available global and regional products or models are utilized. The seasonal and monthly forecasts are derived from the North American Multimodel Ensemble (NMME; Kirtman et al., 2014), whereas the weekly forecast is obtained from a combination of forecasts—NAM (NCEP, 2006), WRF (NCAR, 2019), and ECMWF (ECMWF, 2021)—

assembled by the Peruvian Hydro-Meteorological Agency, SENAMHI, as a consensus prediction. Only the wet season (December-April) is considered. Although these standardized forecast models are useful from an organizational perspective, as they are transparent and globally consistent, their lack of tailoring to local conditions may inhibit their skill and therefore decrease overall benefits (Alexander et al., 2018). This study in part aims to consider how slight differences in forecast methodology may alter optimal decisions, as described in Sections 3-5.

As a case study, we conduct a retrospective analysis, using EAP protocols, from 1982-2018, focusing on the particularly hazardous regions of Piura and Lambayeque (7.5° - 3.5° S, 79.5° - 80.5° W) on the northern coast (Figure 5-2). This study area averages less than 500 mm of precipitation per year, but is prone to extreme rains, which were particularly intense during the extreme El Niño events of 1982-83, 1997-98, and 2017 (Figure 5-3), and which would have likely triggered action had the EAP been in place. Observed precipitation is from the Peruvian Interpolated data of the SENAMHI's Climatological and hydrological Observations (PISCO; Aybar et al., 2019), a gridded daily dataset at 0.1° spatial resolution.

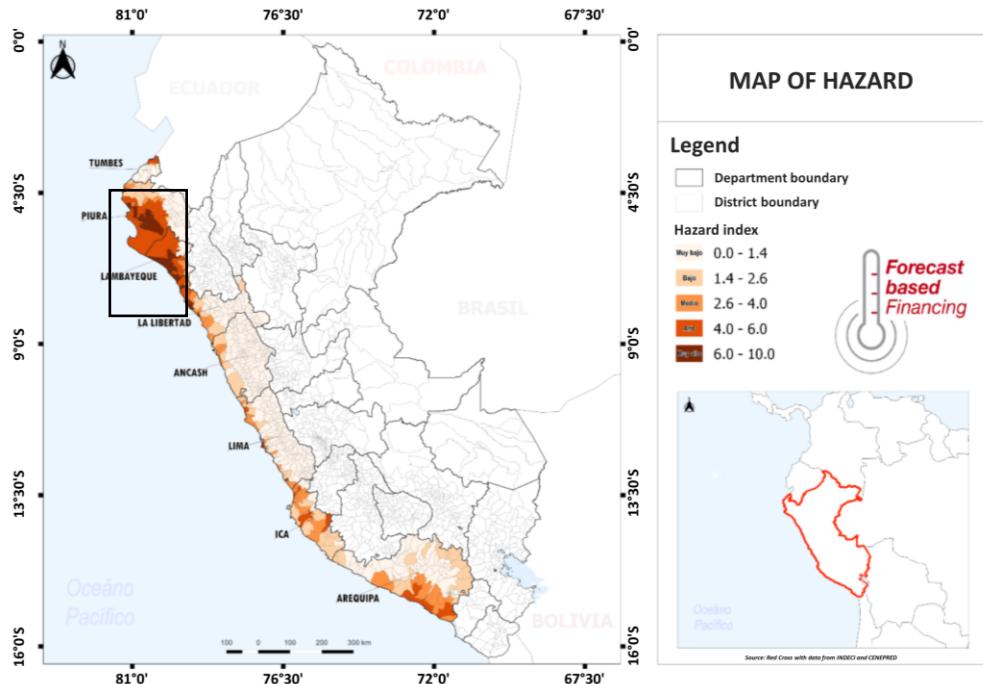


Figure 5-2. Map of El Niño-induced extreme rain hazard by district (Peruvian Red Cross, 2018; based on the methodology of Marin-Ferrer et al., 2017), showing Piura-Lambayeque study region (black box)

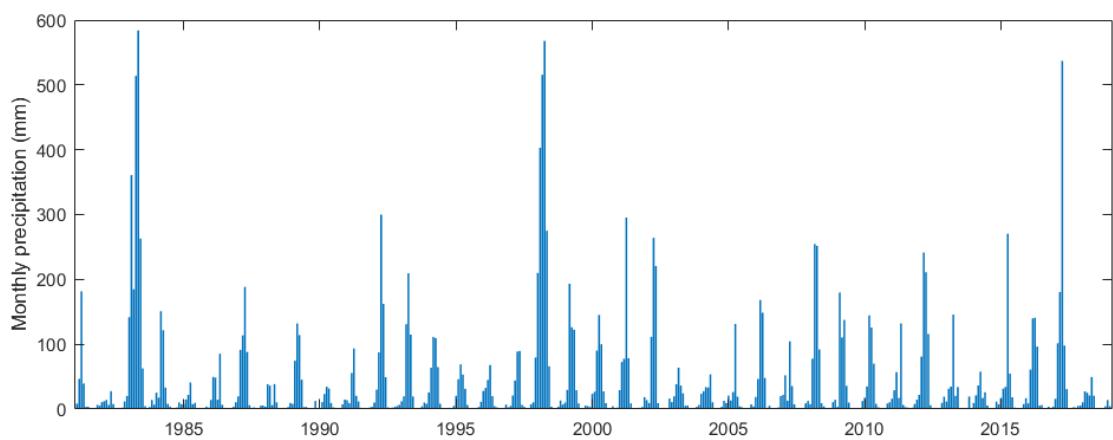


Figure 5-3. Monthly precipitation (PISCO; Aybar et al., 2019) in the Piura-Lambayeque study region

5.3 Modeling framework

5.3.1 Forecasts

To optimize key decisions leading to potential benefits and costs of the EAP, retrospective forecasts are developed. We first prepare seasonal, monthly, and weekly forecasts that the EAP can directly utilize. These forecasts are obtained by performing post-processing on the hindcast outputs of globally available models. We compare two different post-processing methods: partial least squares (PLS; Wold et al., 1984) and logistic regression. Next, we simulate the decisions and actions under the EAP using these forecasts. We measure the performance of these decisions using three performance metrics; two assume risk neutrality (benefit-cost ratio and relative expense reduction) and one assumes risk-averse preferences (CARA utility over net benefits). Finally, we perform a sensitivity analysis using different forecast triggers and examine the effect of forecast methodology (i.e., post-processing methods), cost-loss ratios of early actions, and risk aversion.

The seasonal and monthly retrospective forecasts (hindcasts) are calculated using NMME (Kirtman et al., 2014) model averages from seven constituent models: CanSIPSv2, COLA-RSMAS-CCSM4, GFDL-CM2p1-aer04, GFDL-CM2p5-FLOR-A06, GFDL-CM2p5-FLOR-B01, NASA-GEOS-S2S, and NCEP-CFSv2. The current EAP uses a logistic regression of each model member, from which probabilities of exceedance are averaged to arrive at a final forecast probability. Model members, however, exhibit a wide range of biases and inaccuracies relative to observed precipitation in the region, such that an equal weighting scheme may not produce the most accurate results (Vavrus et al., 2021). To consider the effect of tailored forecasts, a PLS regression of the model members is also employed as an alternative forecasting method. For the PLS regression, the model members are bias corrected via quantile mapping, which has been shown to simply and effectively remove historical biases of climate models relative to observations (Cannon et al., 2015). Next, an ensemble is generated by random bootstrap aggregation from the

set of residuals for each month or season. This PLS forecast thus produces a probabilistic forecast, like the logistic regression framework, but with greater weight on better-performing NMME models. Forecast skill in both cases is measured using a Brier skill score (Brier, 1950) according to the form:

$$BSS = 1 - \frac{BS_{fcst}}{BS_{clim}} \quad (5-1)$$

where BS is the Brier score of the forecast ($fcst$) or of a naïve climatology forecast ($clim$), itself defined as:

$$BS = \frac{1}{N} \sum_{t=1}^N (p_t - o_t)^2 \quad (5-2)$$

where N is the number of time-steps over which an event can occur, p is the forecast probability of exceedance (equivalent to 1 minus the percentile to be exceeded in the case of the naïve climatology forecast; i.e., 0.05 for the seasonal and monthly forecasts and 0.01 for the weekly forecast), and o is the actual outcome (1 for exceedance, 0 otherwise). The seasonal forecasts are issued monthly for December-January-February through March-April-May, while the monthly forecasts are issued monthly for December through April; regressions are calculated individually for each month or three-month period. Weekly forecasts are from the ERA-Interim dataset (Dee et al., 2011). Although technically a reanalysis product, it is preferred as it provides publicly available hindcasts, issued daily, for the entire study period (1982-2018) and also correlates well with the observational PISCO dataset (Imfield et al., 2019). It should be noted, however, that as a reanalysis product, it incorporates some observed variables and thus is likely to be more accurate

than a true reforecast; nevertheless, this study weighs the need to capture the three extreme El Niño events of the last 40 years more heavily than the need to perfectly capture the characteristics of the SENAMHI consensus predictions. The weekly forecasts follow the same post-processing approach as that of the PLS regressions to generate forecast probabilities: bias correction using quantile mapping, followed by random bootstrap aggregation from the set of residuals. Probability distributions and quantiles, however, are calculated over the whole wet season of December-April instead of by month or three-month period. Seasonal and monthly forecasts are developed and issued as single values for the entire study area, whereas weekly forecasts are developed and issued at each 0.1° grid cell, mirroring the EAP's use of longer-lead forecasts for preparation and weekly forecasts for location-specific interventions.

5.3.2 Sensitivity analysis and trigger thresholds

To optimize probabilistic trigger thresholds (τ , see Section 2), a wide range of values varying from 0.01 (near-constant activation) to 0.99 (extreme reluctance to activate) is considered for each forecast. It is assumed that weekly forecast actions cannot be triggered without at least one longer-lead forecast (seasonal or monthly) already activated. Costs and benefits for each combination of actions are estimated from the EAP's itemized budget (Table 5-1; a full budget and calculations used to arrive at benefits is included in supplemental material to Lala et al., 2021b). Generally, seasonal and monthly forecasts trigger personnel coordination and other administrative duties; monthly triggers further include procurement of WaSH supplies and immediate mobilization of personnel to an at-risk region. Since benefits are not explicitly quantified, they are assumed to relate linearly to the costs of beneficial actions (e.g. the cost of medical kits and health volunteers is factored into both benefits and costs, but the cost of office supplies or transportation to the affected community is only factored into costs and not into benefits). To explore the actual value

of benefits, a variable cost - avoidable loss (C/L_a) ratio is employed, which measures the ratio of the total cost of action to the maximum level of avoidable losses (i.e., the full cost of activating all three forecasts divided by the full benefits of taking these actions). For example, a C/L_a ratio of 0.1 would indicate that the cost of full activation—seasonal, monthly, and weekly—is 10% of the value of losses that would be avoided in the event of a disaster. Some actions are low-regret (e.g., sanitation kits can last for several seasons); thus it is assumed that they provide future benefit even if the event does not occur as predicted. We therefore assume a “false alarm” for the weekly forecast actions provide some benefit ($0.63B$, where B is the full benefit of an activation with both seasonal and monthly prepositioning, equivalent to L_a if the extreme event occurs at the time predicted by the forecast) if an extreme event occurs after the intervention but within the same season. Unavoidable losses are calculated as any event within the season that occurs before an intervention; there may be more than one loss event if actions are triggered late or never, incentivizing early action and potentially lower probabilistic trigger thresholds. Since the seasonal and monthly forecasts can only trigger preparatory actions, there are no associated benefits unless the weekly interventions are also triggered.

Table 5-1. Costs and benefits of weekly interventions as a ratio of maximum achievable benefits B and full cost of intervention C . Note that unavoidable losses are not included, as there may be more than one loss event per season.

Event occurred?	Yes	Late	No
Monthly prepositioning triggered?	$0.66B, 0.73C$	$0.42B, 0.73C$	$0.73C$
Seasonal prepositioning triggered?	$0.55B, 0.78C$	$0.35B, 0.78C$	$0.78C$
Seasonal + monthly prepositioning triggered?	B, C	$0.63B, C$	C

Performance metrics for a risk-neutral decision maker take two forms: (1) a benefit-cost ratio, and (2) relative expense reduction. We assume that the decision maker makes early action decisions based on seasonal, monthly, and weekly forecasts, and receives financial returns or utility at the

end of the season once the uncertainty around the loss event resolves, i.e., when the costs, benefits, and losses are realized. Specifically, we define the benefit-cost ratio as:

$$BCR = \frac{1}{n} \sum_{i=1}^n \frac{B_i}{C_i} \quad (5-3)$$

where B_i and C_i are the total benefit and total cost of taking early action based on Table 5-1 in the i th season and n is the total number of seasons in the study period. The second risk-neutral metric, relative expense reduction, is defined as the saved expenses relative to those incurred without using early action (Lopez et al., 2017):

$$V_{rel} = \frac{\sum_{i=1}^n B_i - C_i}{\sum_{i=1}^n L_i} \quad (5-4)$$

where L_i is the total loss incurred due to the loss event in the i th season. The benefit-cost ratio considers the costs of action and avoided losses, whereas relative expense reduction includes losses from missed events in addition to losses from triggered events.

Although a cost-loss framework (i.e., using benefit-cost ratio or relative expense reduction) is useful for measuring actual financial returns, it assumes risk neutrality on behalf of the decision maker, despite evidence that most are risk-averse (Matte et al., 2017). We therefore employ expected utility theory under constant absolute risk aversion (CARA), which aims to more accurately depict the behavior of a disaster risk manager, assuming they behave like a public utility in which wealth effects are negligible (Matte et al., 2017). We thus measure a decision maker's expected utility as:

$$EU = \frac{1}{n} \sum_{i=1}^n u(B_i - C_i - L_i) \quad (5-5)$$

where $u(x) = \frac{-1}{A} e^{-Ax}$ is the agent's utility from the financial outcome x , measured in millions of USD, and A is the Arrow-Pratt risk aversion coefficient (Pratt, 1964; Arrow, 1965), which is strictly positive for a risk-averse decision maker.

Finally, to bridge the concepts of forecast skill and probabilistic threshold optimality, we include the *false alarm ratio* (FAR) as a performance metric. Traditionally, the FAR is defined as the ratio of action taken in vain ("misses") to the total number of actions taken ("hits" plus "misses"). However, given the multi-stage nature of the EAP, we elect to define a *combined* FAR that requires a "hit" at the weekly level (i.e., there is activation of the community intervention and not just preparatory actions); all other actions are counted as misses:

$$FAR_{combined} = \frac{\sum_{S_l} T_l - (H_{s+m+w} + H_{s+w} + H_{m+w})}{\sum_{S_l} T_l} \quad (5-6)$$

where S_l is the set of all combinations of lead time l —seasonal (s), monthly (m), or weekly (w)—for triggered action T , and H is a hit for the given set of lead times. The target community for interventions is selected based on expert opinion and risk assessment; however, for the simplicity and generalization of this study's methodology, final costs, benefits, and utilities are averaged across all grid cells to provide general guidance on financial returns.

5.4 Results

5.4.1 Forecasts

Forecasts at all scales demonstrate bias that is addressed via quantile mapping. Since weekly forecasts are based on a single model, simple bias correction is sufficient. The NMME-based seasonal and monthly forecast models, however, demonstrate bias to varying degrees; the NASA GEOS-S2S and NCEP-CFSv2 models more accurately capture precipitation in northwestern Peru than the others (Figure 5-4). Thus, a weighted PLS regression of members outperforms a pooling approach, although Brier skill scores are generally above zero (i.e., better than a naïve climatology forecast) in both cases (Figure 5-5). Disaggregating the Brier scores (i.e., taking the individual forecast error at each timestep) further illustrates the general trend of capturing extreme El Niño years without generating false alarms in other years (Figure 5-6). While there is some forecast error in extreme years, false alarms in normal years are highly unlikely, as the seasonal and monthly forecasts generally correctly predict a low probability of extreme events and will not activate actions.

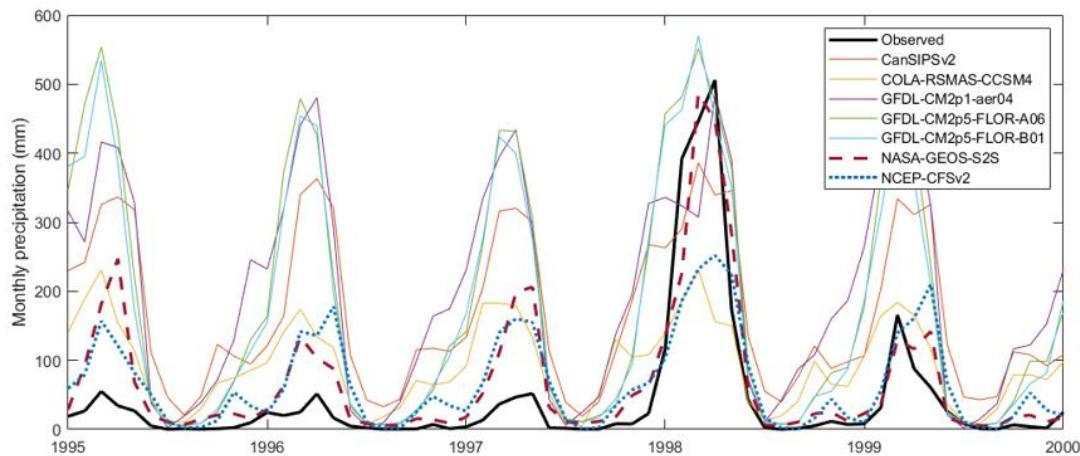


Figure 5-4. Observed monthly precipitation (black line) and NMME modeled precipitation (colored lines) from 1995-2000, showing extreme El Niño event of 1997-98

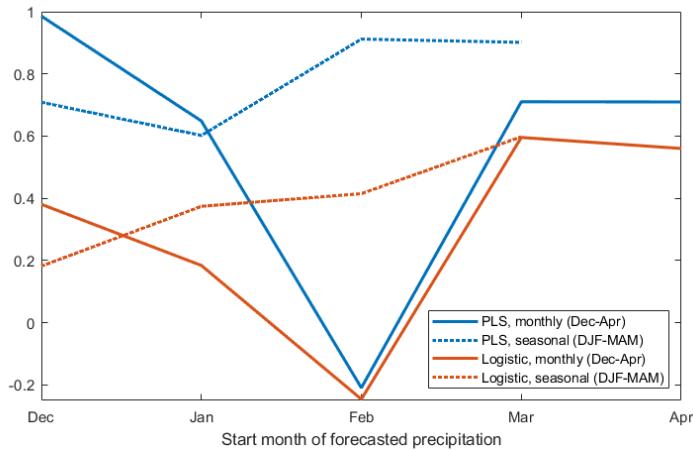


Figure 5-5. Brier skill scores for monthly and seasonal hindcasts over 1982-2018

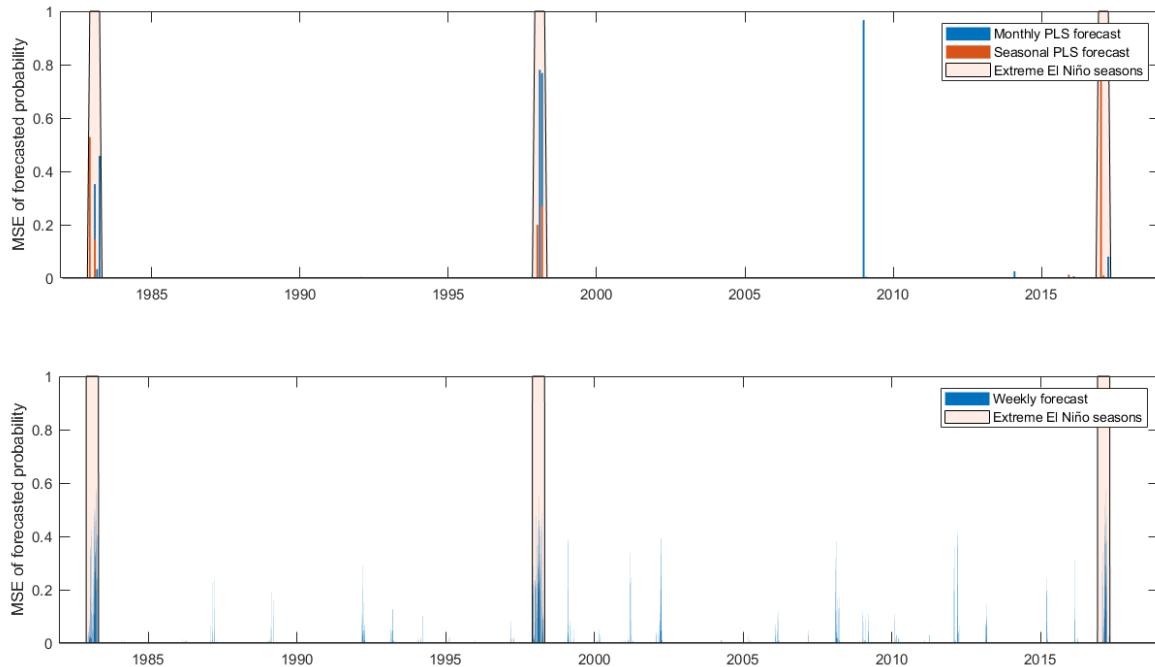


Figure 5-6. Disaggregated Brier scores (i.e., $(p_t - o_t)^2$) for the monthly and seasonal PLS forecasts (top) and weekly forecasts (bottom)

5.4.2 Sensitivity analysis and trigger thresholds

Once forecasts are processed, a sensitivity analysis is performed to identify costs and benefits for various probabilistic trigger thresholds. In general, actions are only triggered during the extreme

El Niño events of 1982-83, 1997-98, and 2017, except for low trigger thresholds. Corresponding benefit-cost ratios are at or near zero in most years, owing to a lack of extreme events and corresponding lack of activations; however, during extreme El Niño years net benefits are generally positive and of high magnitude (Figure 5-7). The use of PLS regression forecasts tends to demonstrate less variance in benefit-cost ratios than does the use of logistic regression forecasts, possibly due to the weighting of more consistently accurate model members. However, benefit-cost ratios based on PLS regression do not necessarily exceed those of the logistic regression forecasts; the greater variance in the latter actually tends to result in higher *maximum* benefit-cost ratios.

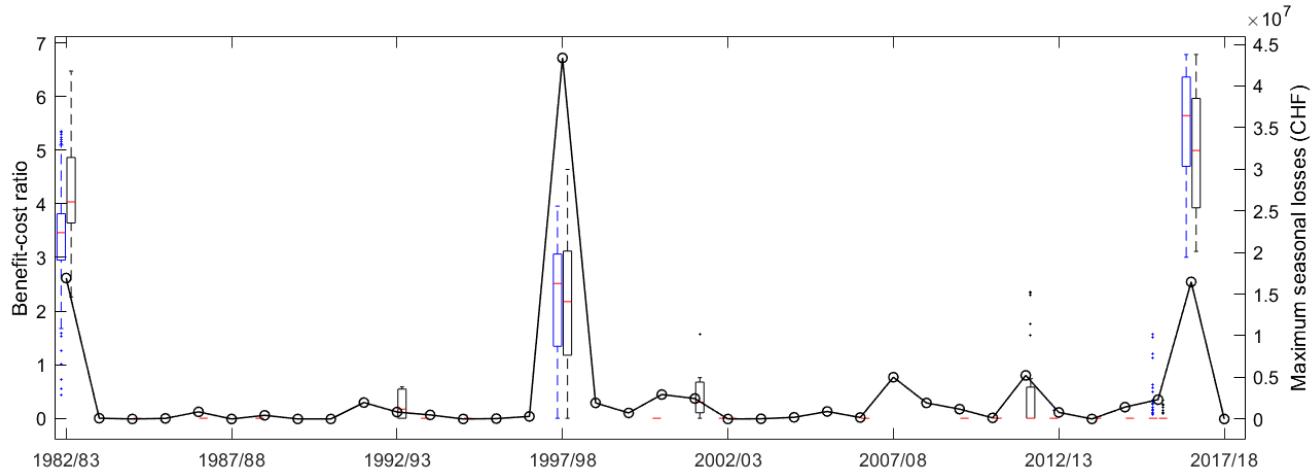
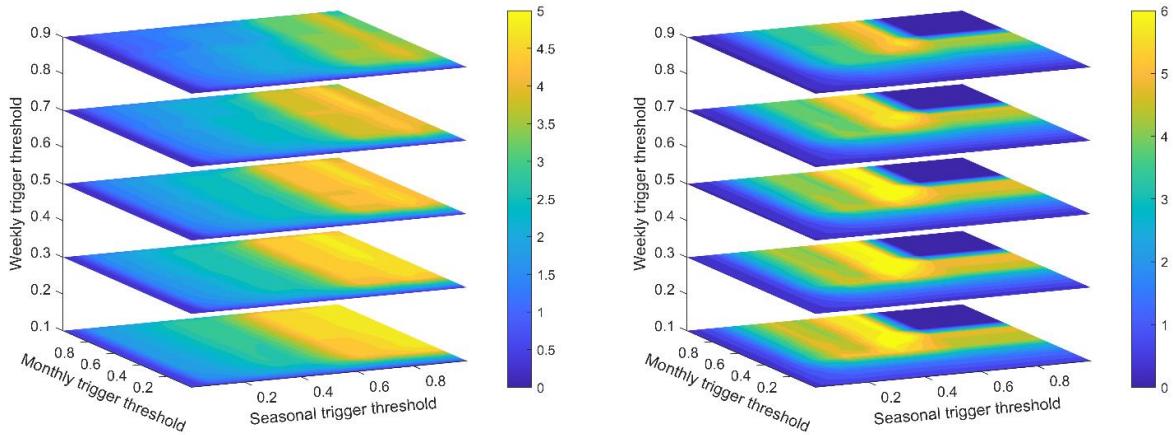


Figure 5-7. Range of possible benefit-cost ratios per season for probabilistic trigger thresholds of 0.1-0.99 and C/L_a ratio of 0.1 for PLS regression (blue boxes) and logistic regression (black boxes). Maximum possible flood-related losses (based on the number of extreme events) illustrated with a black line (right axis)

Overall, optimal monthly and seasonal probabilistic trigger thresholds tend to be higher than optimal weekly probabilistic trigger thresholds (Figure 5-8-Figure 5-9), owing to the conditional nature of weekly interventions (i.e., they cannot be triggered without a seasonal or monthly trigger) and to the extended benefits of most weekly actions, which support communities even if the

predicted extreme event occurs later than expected. Trends in optimality vary mostly with the seasonal and monthly triggers; the weekly trigger mainly influences the magnitude, rather than the relative trend, of benefits. Optimal benefit-cost ratios tend to have slightly higher probabilistic thresholds than do optimal relative expense reductions (Figure 5-8). Notably, the increased skill of the PLS forecasts does not translate into higher optimal performance metrics, although metrics are more consistent throughout the decision space relative to the logistic regression forecasts. In optimizing utility, the C/L_a ratio is notably more influential than the risk aversion coefficient A . An increase in the C/L_a ratio tends to increase optimal probabilistic trigger thresholds at all scales, whereas A has minimal impact on optimal trigger thresholds (Figure 5-9). Depending on the C/L_a ratio and level of A , this increase in utility corresponds to approximately \$300,000-\$2.2 million in financial savings. We note, however, that the levels of A used are rather high; realistic values would likely be lower, although this does not substantially affect the results in terms of optimality. Also of note is the logistic regression forecasts' lack of financial returns at higher trigger thresholds (τ_s and $\tau_m \geq 0.7$), due to a failure to trigger any actions above these levels.



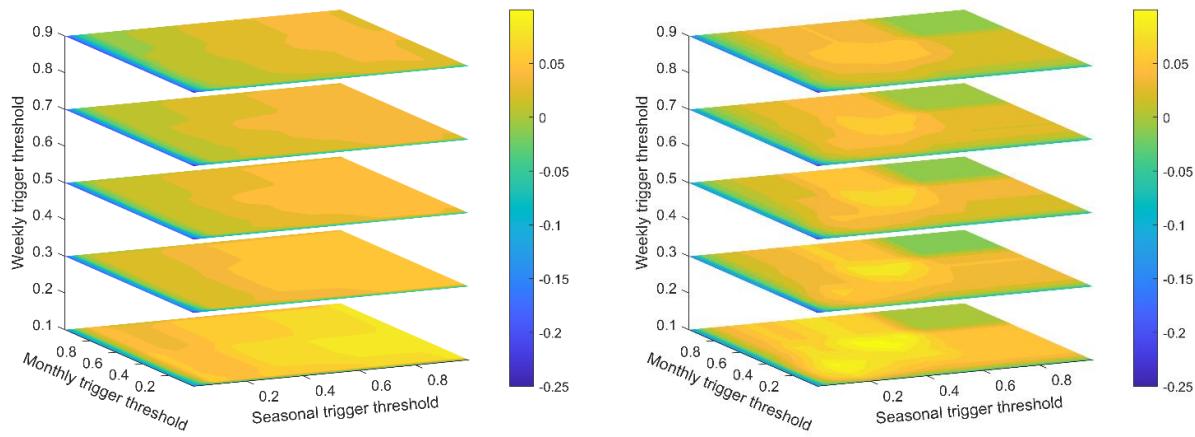


Figure 5-8. Benefit-cost ratio (top) and relative expense reduction (bottom) averaged over 1982-2018 for various probabilistic trigger thresholds and a C/L_a ratio of 0.1, using PLS forecasts (left) and logistic regression forecasts (right)

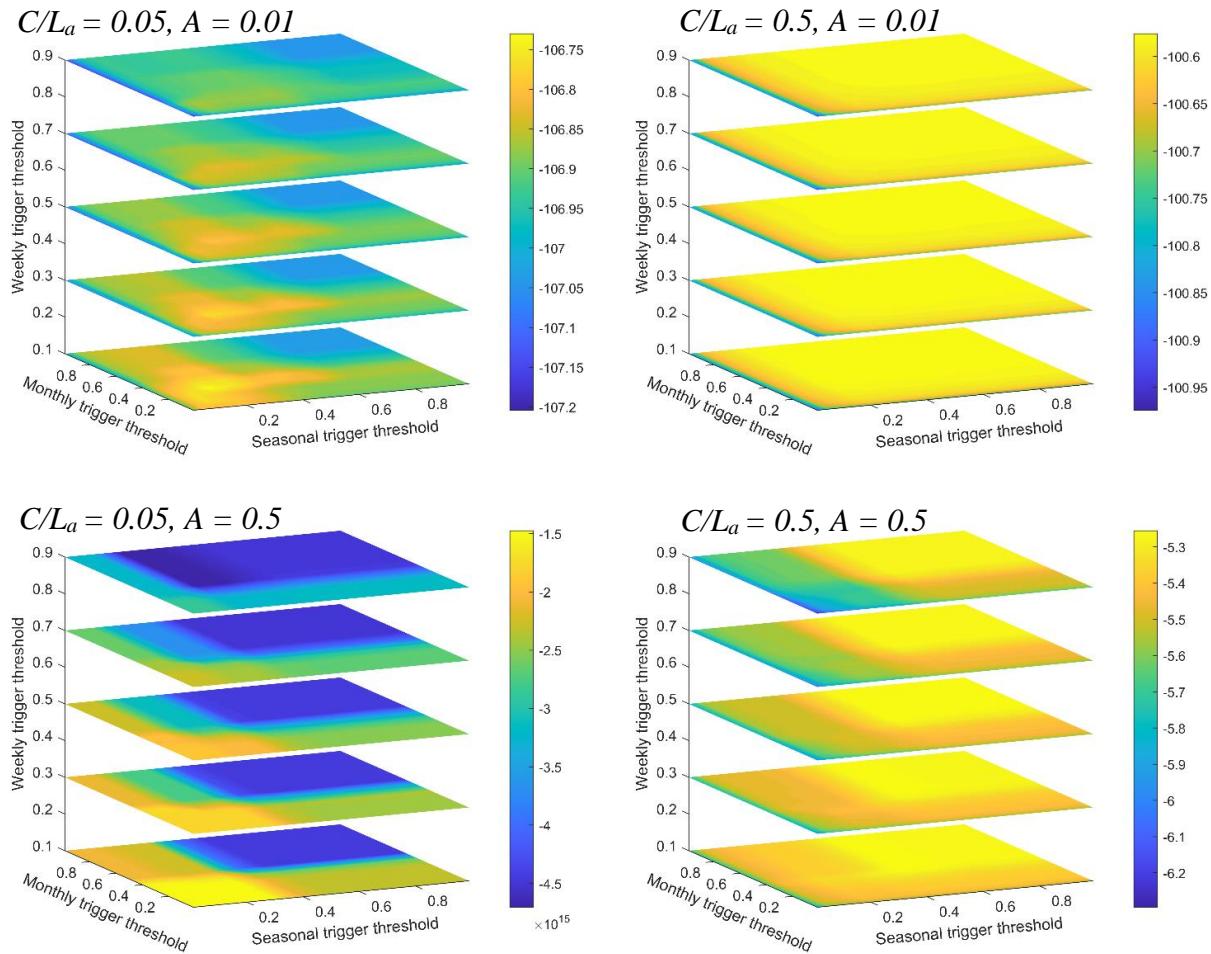


Figure 5-9. Utility with $A = 0.01$ (top) and $A = 0.5$ (bottom), for a C/L_a ratio of 0.05 (left) and 0.5 (right), using logistic regression forecasts

The combined false alarm ratio follows a very similar pattern to that of the benefit-cost ratio, with moderately high seasonal and monthly probabilistic trigger thresholds and moderately low weekly probabilistic trigger thresholds being optimal (Figure 5-10). In general, FARs are relatively high, given the strict definition used capturing multiple scales (e.g., a “hit” at the seasonal or monthly scale that is not coupled with a “hit” at the weekly scale is defined as a false alarm for the purposes of this study), but optimal FARs still fall at or below 0.5, which may avoid risks associated with the “false alarm effect” in which frequent false alarms hinder the sustainability of early warning systems (Barnes et al., 2006).

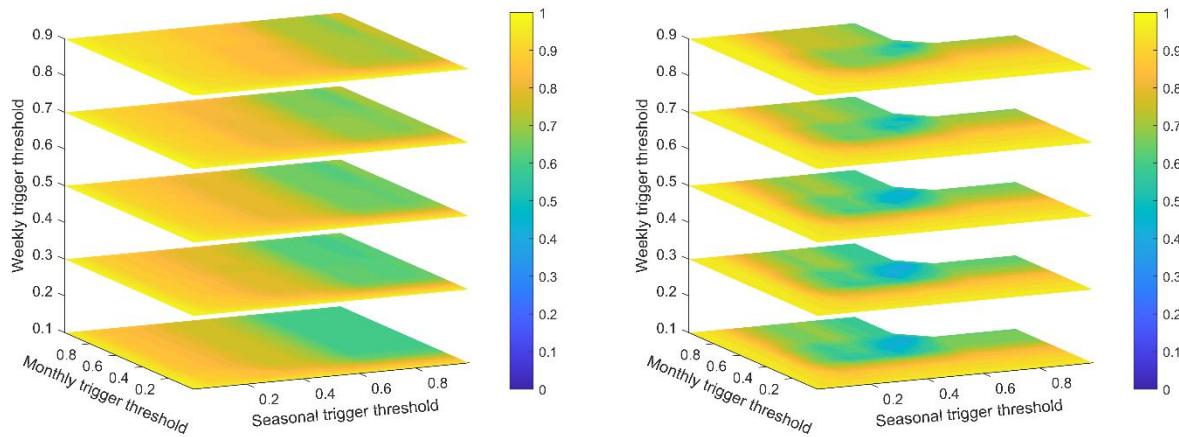


Figure 5-10. Combined false alarm ratio averaged over 1982-2018 for various probabilistic trigger thresholds, using PLS forecasts (left) and logistic regression forecasts (right)

Finally, it should be noted that overall project costs exceed benefits for a C/L_a ratio greater than ~0.5 for the PLS forecast and ~0.7 for the logistic forecasts (Figure 5-11). This suggests that avoided losses from an early intervention should be valued at 1.4 – 2 times the project cost, although these need not be material losses. Avoided deaths or injuries, for example, may be valued well above these numbers, making the project feasible, even if material losses amount to less than the break-even monetary value.

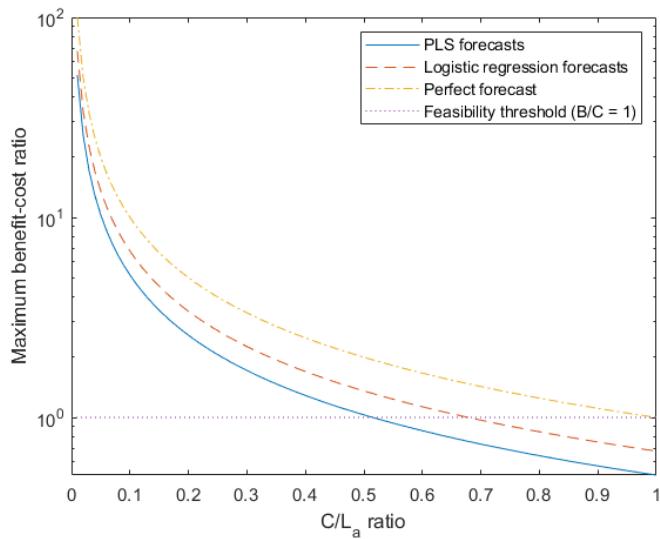


Figure 5-11. Maximum benefit-cost ratio as a function of C/L_a ratio for the two forecast methods, a perfect forecast (the inverse of the C/L_a ratio) and a feasibility threshold ($B/C = 1$)

5.5 Discussion and Conclusions

This study presents a framework for optimization of operational early action protocols, with a demonstration for extreme precipitation in coastal Peru. Results indicate that (1) overall, moderate-to-high seasonal and monthly probabilistic trigger thresholds along with relatively lower weekly trigger thresholds are optimal, and thus preparatory actions at longer leads are warranted; (2) the cost-to-avoided loss ratio C/L_a has primary influence over optimal trigger thresholds; and (3) increased forecast skill does not necessarily increase optimal returns.

In general, the conservative approach to preparatory actions (i.e., moderate-to-high seasonal and monthly triggers) reduces the risks of acting in vain, particularly since there is no benefit associated with actions at the seasonal and monthly lead. Moreover, by ensuring confidence in an extremely wet season or month, this conservative preparatory strategy allows for a more liberal approach at the weekly scale, when actions may provide some benefit even if taken several weeks before an extreme rainfall event occurs.

Secondly, while forecast skill and the choice of performance metrics play a role in optimal probabilistic trigger thresholds, the relative magnitude of performance metrics (particularly C/L_a) is paramount in determining optimal thresholds. Given the relatively lower cost yet more beneficial actions at the monthly level relative to the seasonal level, an optimal program would require a slightly lower probabilistic threshold for the monthly forecast than for the seasonal forecast. Optimal weekly forecast thresholds would be even lower, since their low-regret nature provides benefits even if the predicted extreme event arrives late. Overall, the differences in optimal probabilistic thresholds at different lead times suggests that there is considerable value in using seasonal and subseasonal forecasts for preparation.

While much of the relative value of the monthly forecast is due to the nature of the EAP's prepositioning protocol, part may also be attributable to the 2017 El Niño rains which fell almost entirely within the month of March; this was better predicted with the monthly—but not seasonal—forecasts (see Figure 5-6). More generally, the differences in temporal scale for each forecast lead play a role in plan effectiveness; a wet season is not necessarily a good indicator of extreme events at the daily or weekly scale, whereas a wet month is more likely to contain shorter periods of extreme rainfall. Despite this lack of direct correlation between the weekly forecast and the longer lead (monthly or seasonal) forecasts, we assert that the benefits of extra lead time by using the longer lead forecasts outweigh the costs of inflexibility in requiring them before triggering the weekly interventions.

Thirdly, in this study, the PLS regression forecasts outperform the logistic regression forecasts in skill, but the greater overall variance in benefits and costs given by the logistic forecasts results in generally higher *optimal* benefit-cost ratios, relative expense reductions, and utilities. This suggests that forecast accuracy is not primary in determining singular optimal probabilistic trigger

thresholds for an EAP. The main benefit of the PLS forecasts is therefore in consistency; variance in the performance metrics is lower and triggers demonstrate more consistent effects on performance metrics, whereas the logistic forecasts abruptly stop triggering at probabilities above 0.7 (see Figure 5-8). Furthermore, the relative accuracy of certain model members—specifically, NASA-GEOS-S2S and NCEP-CFSv2—in northwestern Peru is not generalizable globally. Moreover, while the use of quantile mapping for bias correction is useful for retrospective studies, future analyses that include climate projections would likely require other methods that account for long-term changes (Cannon et al., 2015). Thus, the relative value of tailored forecasts—if any—must be evaluated for each project, requiring that humanitarian actors balance customized methodologies with operational efficiency and considerations of optimality.

The work presented here fills several research gaps on forecast valuation, the role of seasonal and subseasonal forecasts, and the implementation of projects within a planetary health perspective. Regarding forecast valuation, this study explicitly considers attitudes towards risk without presenting hardships to beneficiaries; risk is borne by the relief agency rather than a marginalized community, and risk aversion is explicitly assessed through expected utility theory. This study also investigates the role of spatiotemporal scales in forecast value, by considering forecasts at two spatial and three temporal scales, which are often ignored in valuation studies (Soares et al., 2018). Finally, the approach is co-evaluated with the end user—the Red Cross—which serves to both clarify methodological issues and encourage the uptake and use of information. Regarding the role of subseasonal forecasts, this project also furthers the goals of the Subseasonal-to-Seasonal (S2S) Prediction Project, which aims to enhance “operational infrastructures and user applications” using forecasts at monthly and seasonal scales (WMO, 2018). In the context of the EAP, subseasonal (i.e., monthly) forecasts may be particularly valuable in bridging the gap between longer-term

ENSO forecasts and individual extreme weather events, given their relative value in the optimization of probabilistic trigger thresholds. Finally, risk reduction measures are considered within the context of the newly emerging field of planetary health by exploring adaptation to climate variability using management decisions over physical infrastructure—i.e., without unsustainably impacting the natural environment (Horton and Lo, 2015).

To the authors' knowledge, this study is the first to incorporate expected utility theory in an FbF framework, providing a more realistic perspective on decision-making. While the use of utility as a performance metric brings insight to operational plans relative to risk-neutral assumptions, the modelling framework presented here is still hindered by some real-world constraints. For example, at low levels of C/L_a , utility tends to be higher for very low probabilistic trigger thresholds, given that a risk-averse decision maker would likely be willing to bear the cost of near-constant activation in order to reduce losses in years with extreme events, even if overall benefit-cost ratios are lower. In reality, however, budget constraints are likely to limit frequent activation and payout, prompting higher trigger thresholds (the Red Cross, for example, restricts payouts to a minimum 5-year return period). The model represented here tends to only activate in extreme El Niño years (i.e., once every 15-20 years), well within Red Cross limits, but is still subject to occasional false alarms for lower trigger thresholds. To capture budget constraints more accurately, this study restricted weekly interventions to once per year but made no constraints on seasonal or monthly actions. Given no such constraint, it is likely that more actions would be triggered in extreme El Niño years, as there tend to be multiple instances of extreme rainfall at the weekly scale during these periods. The outcome of the Red Cross' 2015/16 pilot project likewise incentivizes higher trigger thresholds, as actions were triggered based on a predicted extreme coastal El Niño that never materialized (Peruvian Red Cross, 2016; Bazo et al., 2019).

This study provides a realistic application of FbF by evaluating an operational EAP; however, limitations still hinder a full portrayal of on-the-ground actions. Budget constraints, as mentioned above, are not explicitly considered other than a maximum of one weekly intervention per year. Additionally, while the cost of acting in vain in terms of reduced financial benefits is included, there is no quantitative consideration of the behavioral “false alarm effect” in terms of program effectiveness, due to mixed characterization of its magnitude in literature (e.g., Barnes et al., 2006; Trainor et al., 2015). Moreover, this study assumes relatively low levels of regret for weekly actions, incentivizing lower probabilistic trigger thresholds. More realistic trigger thresholds may thus be slightly higher, especially for the weekly forecast, which is optimally triggered at probabilities below 50% under current model assumptions. The use of a variable C/L_a ratio allows for flexibility in defining benefits, but immaterial losses, including death or injury, remain diffusely defined (Yu and Tang, 2017; Huang et al., 2018). Finally, extreme events at the seasonal and monthly scale are defined by the 95th percentile for the month or season predicted which may lead to vastly different rainfall totals for a given time period (for example, the 95th percentile of cumulative monthly rainfall in northwestern Peru ranges from approximately 100 mm in December to nearly 530 mm in March). Although a single value over the full wet season is used for the weekly forecast, the monthly and seasonal forecasts may lack the sharpness to properly characterize catastrophic rainfall-induced floods.

Given the conclusions—and limitations—of this study, we propose a few ideas for future work. Firstly, this study does not explicitly consider future climate change, including the likelihood of more frequent extreme events (Khalil et al., 2007; Reguero et al., 2015). This was primarily motivated by the availability of observed data and forecasts for a retrospective analysis; however, characterizing the future climate of coastal Peru in subsequent FbF modeling studies, particularly

those emphasizing theoretical approaches, may be warranted. Secondly, a stronger focus on community impacts—particularly the long-term effects of working within a region for an extended period—would likely increase the effectiveness of future FbF programs. Evidence has shown that even relatively short periods of consistent aid can have long-lasting positive effects (Butler, 2015), but much remains to be seen given the novelty of FbF. Finally, we suggest that other FbF programs conduct similar sensitivity analyses, both to avoid the pitfalls of the 2015-16 pilot study (Bazo et al., 2019) and given the tendency for forecast skill and costs and benefits to be highly case-specific.

Chapter 6. Evaluating prospects for subseasonal-to-seasonal forecast-based anticipatory action from a global perspective

Adapted from: Lala, J., Lee, D., Bazo, J., and Block, P. Chapter 1. Evaluating prospects for subseasonal-to-seasonal forecast-based anticipatory action from a global perspective. In draft.

6.1 Introduction

Extreme climate and weather events are a major impediment to global development and are expected to become more frequent as climate changes (IPCC, 2014). The subsequent push for risk reduction measures has prompted relief agencies and governments to consider management options in addition to physical infrastructure, given the potential adverse effects of the latter and the relatively low capital costs of the former (Di Baldassarre et al., 2018; Kelman, 2013). A more recent innovation in this field is forecast-based anticipatory action, in which forecasts are used to trigger action before a disaster occurs. Much of the impetus for the rise of early action protocols (EAPs) results from gains in the field of subseasonal-to-seasonal (S2S) forecasts, either alone or as complements to shorter-lead forecasts (Vitart and Robertson, 2018). Indeed, in-depth studies of EAPs have found that these medium-term forecasts, ranging from months to seasons in lead time, can have a major impact on project efficacy and cost as part of a “Ready-Set-Go!” approach (Bazo et al., 2019; Lala et al., 2021b). Forecast systems used for EAP development are independently developed by various agencies and organizations, such as the National Centers for Environmental Predictions (NCEP) Climate Forecast System version 2 (CFSv2; Saha et al., 2014) and the European Centre for Medium-Range Weather Forecasts (ECMWF) Seasonal Forecast System 5 (SEAS5; Johnson et al., 2019). In most cases, these systems provide a forecast ensemble based on various initial conditions and assumptions, allowing for probabilistic analysis that generally

enhances user applications, particularly for longer lead times (Matte et al., 2017). The improvement in skill and spatial resolution of forecasts coupled with innovations in humanitarian aid has thus provided a clear roadmap for the development of EAPs and anticipatory action in general.

EAPs have been developed by the Red Cross and other aid organizations in Latin America, East and West Africa, and South and Southeast Asia (Coughlan de Perez et al., 2015; Lopez et al., 2017; IFRC, 2020). Despite the widening geographic scope of these projects, there remain outstanding questions on appropriately targeting high-risk populations. As the global community shifts from disaster management to disaster *risk* management (United Nations, 2015), there is an increasing need to identify risk as a distinct component of disaster response. The European Commission has launched the Index for Risk Management (INFORM) project to catalogue and share physical and socioeconomic factors that constitute disaster risk (Marin-Ferrer et al., 2017). Data is available for 191 countries at the national scale, and at the subnational scale for a few select locations, allowing for identification of risk based on a variety of hydro-meteorological hazards, including floods and droughts. Although this data is useful for identifying socio-economic and population risk, there is minimal integration with early warning systems to properly anticipate and mitigate disasters (OCHA, 2021). The identification of open, skillful, and global forecasts and standardized risk metrics is thus a major next step in the implementation of sustainable anticipatory action programs.

This study addresses outstanding challenges in disaster risk reduction by integrating global S2S precipitation forecasts with a standard, global definition of disaster risk. As many global disaster risk models are limited by granular data (Ward et al., 2015), we incorporate, where available, subnational data on risk, while also downscaling and bias correcting forecasts using gauge-corrected precipitation datasets to ensure a spatial scale fine enough for targeted relief efforts. At

the country level, we conduct a global analysis, while at the subnational level, we focus on the Greater Horn of Africa, which is among the most drought-vulnerable regions in the world (Ahmadalipour and Moradkhani, 2018), as well as Peru, which is highly vulnerable to floods and extreme rainfall (Bazo et al., 2019).

6.2 Methodology

This study consists of three parts: (1) quantifying and evaluating risk scores for floods and droughts, (2) evaluating global S2S precipitation forecast skill through bias correction and spatial disaggregation, and (3) integrating the two into an anticipatory action “suitability score” in which areas with coincident high predictability skill and high risk scores are highlighted.

6.2.1 Evaluation of disaster risk

Risk data at the national level, as well as the regional level for the Greater Horn of Africa, comes from the Index for Risk Management (INFORM) project (Marin-Ferrer et al., 2017). INFORM defines risk as the product of three separate components: (1) hazard and exposure, including the existence of extreme events and populations exposed to them; (2) vulnerability, comprising of socioeconomic health factors that leave communities susceptible to a hazard; and (3) lack of coping capacity, indicating a lack of available resources to lessen the impact of a hazard. Each component is broken down into multiple subcomponents, collectively comprising 32 indicators, and normalized to a score of 0-10 (Figure 6-1).

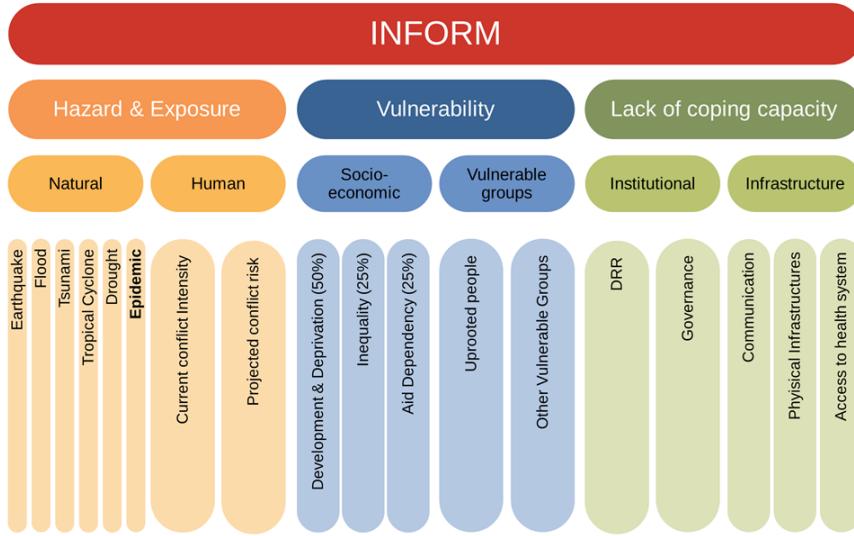


Figure 6-1. Components of INFORM risk index (source: INFORM, 2020)

Given these components, INFORM then calculates a risk score, ranging from 0 to 10, as follows:

$$Risk = Hazard \& exposure^{\frac{1}{3}} \times Vulnerability^{\frac{1}{3}} \times Lack \& coping \& capacity^{\frac{1}{3}} \quad (6-1)$$

with a score of 0 – 1.9 indicating very low risk, 2.0 – 3.4 low risk, 3.5 – 4.9 medium risk, 5.0 – 6.4 high risk, and 6.5 – 10 very high risk. Given that we investigate only precipitation hazards, the natural component of hazard and exposure is restricted to either flood or drought depending on the forecast; indicators relating to earthquakes, tsunamis, cyclones, or epidemics are disregarded in this analysis.

Although INFORM provides this data at the national level for 191 countries, disaster risk models are often limited by the coarse resolution of their input data (Ward et al., 2015). INFORM has addressed this by issuing subnational data for some regions, including parts of sub-Saharan Africa, Central America, and the eastern Mediterranean; however, subnational data in most of the world is sparse. We therefore consider two case studies for subnational data: the Greater Horn of Africa at the regional/provincial (Level 1) scale, using INFORM data, and Peru at the district (Level 2)

scale, using data gathered from the Peruvian government and the Red Cross. The Greater Horn of Africa is highly drought-vulnerable, and risk of drought is only expected to increase in the region due to climate change (Ahmadalipour and Moradkhani, 2018). Contrastingly, Peru is a data-rich country that is highly susceptible to floods and extreme rainfall, particularly in the coastal north and Amazonia (Bazo et al., 2019).

Given the different sources of data for Peru, an exact one-to-one correspondence with INFORM indicators is difficult; however, we follow the same general approach of INFORM as the product of hazard and exposure, vulnerability, and lack of coping capacity. Lee et al. (2021) found that an equal weighting of socio-health vulnerability and coping capacity indicators was best correlated with forecast-informed estimates of flood impact; thus, we combine these socioeconomic indicators into a combined vulnerability and lack of coping capacity score with equal weighting of all indicators:

$$Risk = Hazard \& exposure^{\frac{1}{3}} \times Vulnerability \& Lack\ of\ coping\ capacity^{\frac{2}{3}} \quad (6-2)$$

To account for the potential differences in risk arising from a different set of metrics, the final risk score is normalized such that the average score over all districts equals the country-level score given by INFORM. All vulnerability and lack of coping capacity are derived from the Peruvian National Institute for Statistics and Informatics (INEI – Instituto Nacional de Estadística e Informática; INEI, 2017) and from the Netherlands Red Cross 510 Dashboard (Netherlands Red Cross, 2021). Hazard and exposure data for floods is the average of three normalized variables: number of damaged houses (INEI), number of affected people (INEI), and percent of area at risk of inundation from maximum daily precipitation (Aybar, 2018). For droughts, hazard and exposure data are derived from the water scarcity analysis of Veldkamp et al. (2016), aggregated to the

regional scale via the ThinkHazard! database (GFDRR Labs, 2021). Indicators in all cases are normalized via minmax normalization; data that is not distributed normally is first transformed via a Box-Cox power transformation (Box and Cox, 1964).

6.2.2 Evaluation of forecasts

Before evaluating forecast skill, a quasi-global timeseries of monthly precipitation for peak months and seasons are determined using CHIRPS, a satellite-based, gauge-corrected, quasi-global (50°S-50°N) precipitation dataset provided at 0.05° spatial resolution. Given that wet season totals tend to be particularly important for agriculture and runoff generation (Funk et al., 2019), precipitation and forecasts are only considered for the peak month—that is, the month in each grid cell with the highest mean climatological value—and peak season, which is defined as the peak month and one month to either side (e.g., if the highest mean precipitation in a location occurs in December, the peak season would be defined as November-December-January). Although the season with the highest mean precipitation may not necessarily correspond to the month with the highest mean, we nevertheless use this definition in order to conserve the “Ready-Set-Go!” approach of the Red Cross EAPs such that action at the seasonal and monthly scale can be activated concurrently. Figure 6-2 depicts the peak month of each grid cell in the dataset (note that CHIRPS only extends from 50°S-50°N, thus parts of North America, Europe, and Asia are not considered in this analysis).

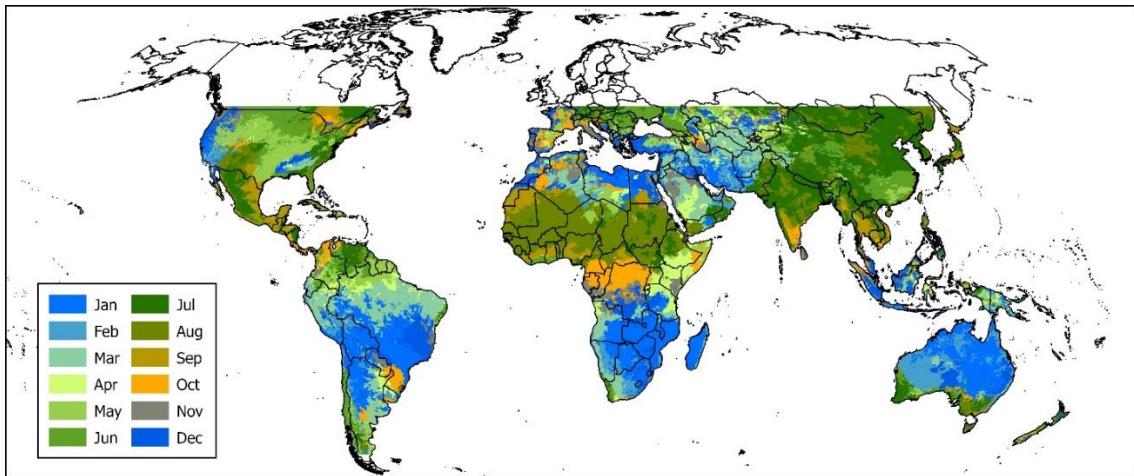


Figure 6-2. Peak month of precipitation, based on highest mean, 1981-2020

Two seasonal-to-subseasonal forecast products are used for the global analysis: the National Centers for Environmental Predictions (NCEP) Climate Forecast System version 2 (CFSv2; Saha et al., 2014), and the European Centre for Medium-Range Weather Forecasts (ECMWF) Seasonal Forecast System 5 (SEAS5; Johnson et al., 2019). Both centers issue monthly forecasts with six (ECMWF) and nine (NCEP) months lead time, and consist of 24 (NCEP) and 25 (ECMWF) ensemble members extending over their available records. Forecasts are evaluated for the peak month (1 month lead) or season (1-3 month lead) at each CHIRPS grid cell over the period 1982-2020 (NCEP) and 1993-2020 (ECMWF) and are spatially disaggregated and bias corrected approximately following the method of Lorenz et al. (2021). First, the individual ensemble members of the forecasts are downscaled to the 0.05° resolution of the precipitation dataset using bilinear interpolation, then, the disaggregated forecasts are bias corrected using quantile mapping of the empirical CDFs of the forecasts to the empirical CDFs of the precipitation (Cannon et al., 2015). Empirical CDFs are chosen as they tend to perform better in bias correction than do parametric CDFs (Gudmundsson et al., 2012). Although individual ensemble members are often

pooled before performing bias correction, we find that bias correction of ensemble members individually generally results in better performance, and we elected to use the latter approach.

Defining floods and droughts is nontrivial; however, past EAP projects provide insight into coherent definitions. The Red Cross's current EAP trigger for extreme precipitation in Peru uses the 95th percentile of monthly or seasonal precipitation in its definition (Bazo et al., 2019; Lala et al., 2021b), while a project in England and Wales used a 1-in-30 year return period for rainfall intensity (~97th percentile; Coughlan de Perez et al., 2015). Yet another project for floods used the 2- to 20-year flood return period (50th-95th percentile; Bischiniotis et al., 2020). Given these studies, we select the 95th percentile of monthly or seasonal rainfall as the threshold for flood-based EAPs. Ideally, riverine flooding based on flow rates or river stage may be preferable, however we select precipitation based on limited riverine data observations and coarse resolution in global hydrologic models (Ward et al., 2020). Regarding droughts, past studies consider the bottom 15%-40% of crop yields (Guimarães Nobre et al., 2019); however, given our use of meteorological—rather than agricultural—drought, we elect to define drought as the 15th percentile of precipitation, based on the definition of “moderate drought” by the U.S. Drought Monitor as the 10th to 20th percentile of precipitation (NDMC, 2021).

Forecast skill is evaluated using the Brier skill score (Brier, 1950) according to the form:

$$BSS = 1 - \frac{BS_{fcst}}{BS_{clim}} \quad (6-3)$$

where BS is the Brier score of the forecast ($fcst$) or of a naïve climatology forecast ($clim$), defined as:

$$BS = \frac{1}{N} \sum_{t=1}^N (p_t - o_t)^2 \quad (6-4)$$

where N is the number of years in the study period, p is the forecast probability of exceedance (defined as the proportion of ensemble members predicting exceedance; i.e., 5% for floods and 15% for droughts), and o is the actual outcome (1 for exceedance, 0 otherwise). Forecasts thus have a BSS between $-\infty$ and 1, with 1 representing a perfect forecast and values greater than 0 indicating an improvement over the naïve climatology forecast.

6.2.3 Identifying suitability scores

Finally, once a risk score (ranging from 0 to 10) and a Brier skill score (ranging from $-\infty$ to 1) are determined for a given location, a *suitability* score is determined as the product of the two:

$$Suitability = Risk * BSS_{max} \quad (6-5)$$

where BSS_{max} is the maximum BSS value, between the NCEP or ECMWF model, at a given grid cell. Suitability can range from $-\infty$ and 10, with 10 indicating a perfect score (i.e., maximum risk and perfect predictability), and any score above 0 indicating some degree of positive forecast skill and risk score.

6.3 Results

6.3.1 Forecasts

Overall, forecast skill is generally modest, with most skill scores in the low positive ranges (Figure 6-3). Skill is spatially heterogeneous, although some areas demonstrate relatively high skill. Moreover, when using the best of either the NCEP or ECMWF forecasts, very few locations have skill scores below zero. In general, we attribute the modest skill to the focus on extreme events

within a relatively short study period (1982-2020 for NCEP, 1993-2020 for ECMWF); though not reported in this study, skill scores for more moderate percentiles (e.g., median precipitation) tend to be higher. For the drought (15th percentile) forecasts, parts of eastern Brazil, East Africa, Afghanistan, and Indonesia have skill scores near 0.5, while for floods (95th percentile), coastal Ecuador and Peru, the Horn of Africa, and parts of Ukraine, Mongolia, and China perform well.

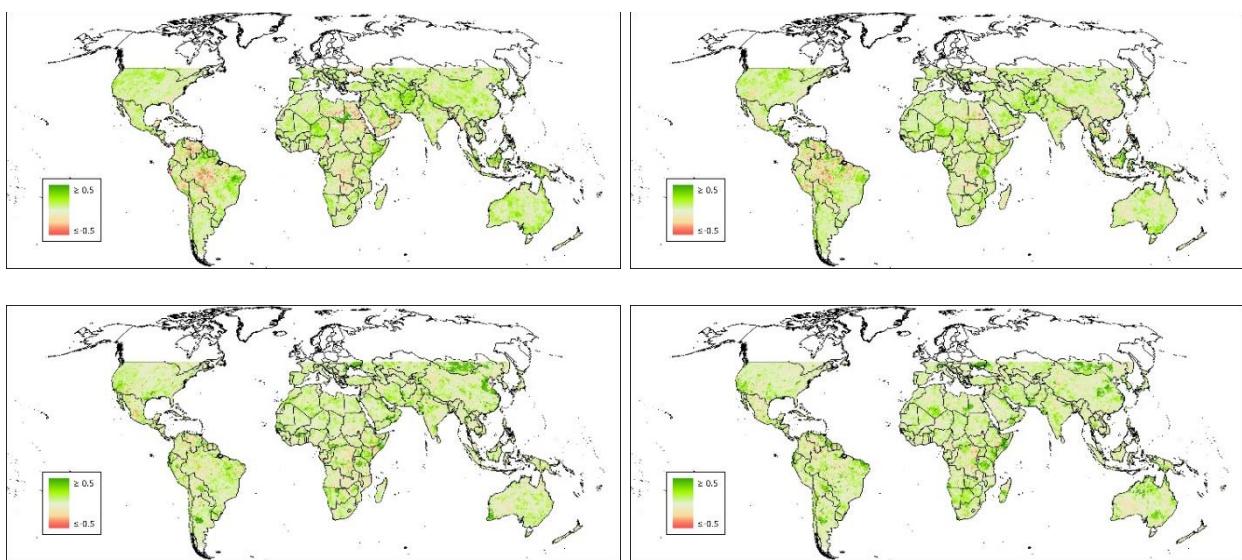


Figure 6-3. Maximum Brier skill scores for NCEP and ECMWF forecasts (BSS_{max}), for peak month (left) and season (right), and for the 15th percentile (top; droughts) and 95th percentile (bottom; floods) of monthly precipitation

6.3.2 Risk and suitability

Globally, risk for both droughts and floods is highest in Africa and conflict-ridden countries of Asia (Figure 6-4, top row). Correlation between drought and flood risk is high due to equivalent vulnerability and lack of coping capacity metrics; only hazard differs between the two. Suitability scores approximately mirror Brier skill scores but with an amplification effect in lower income countries—including much of Africa—and attenuation in wealthy countries (Figure 6-4, middle and bottom rows).

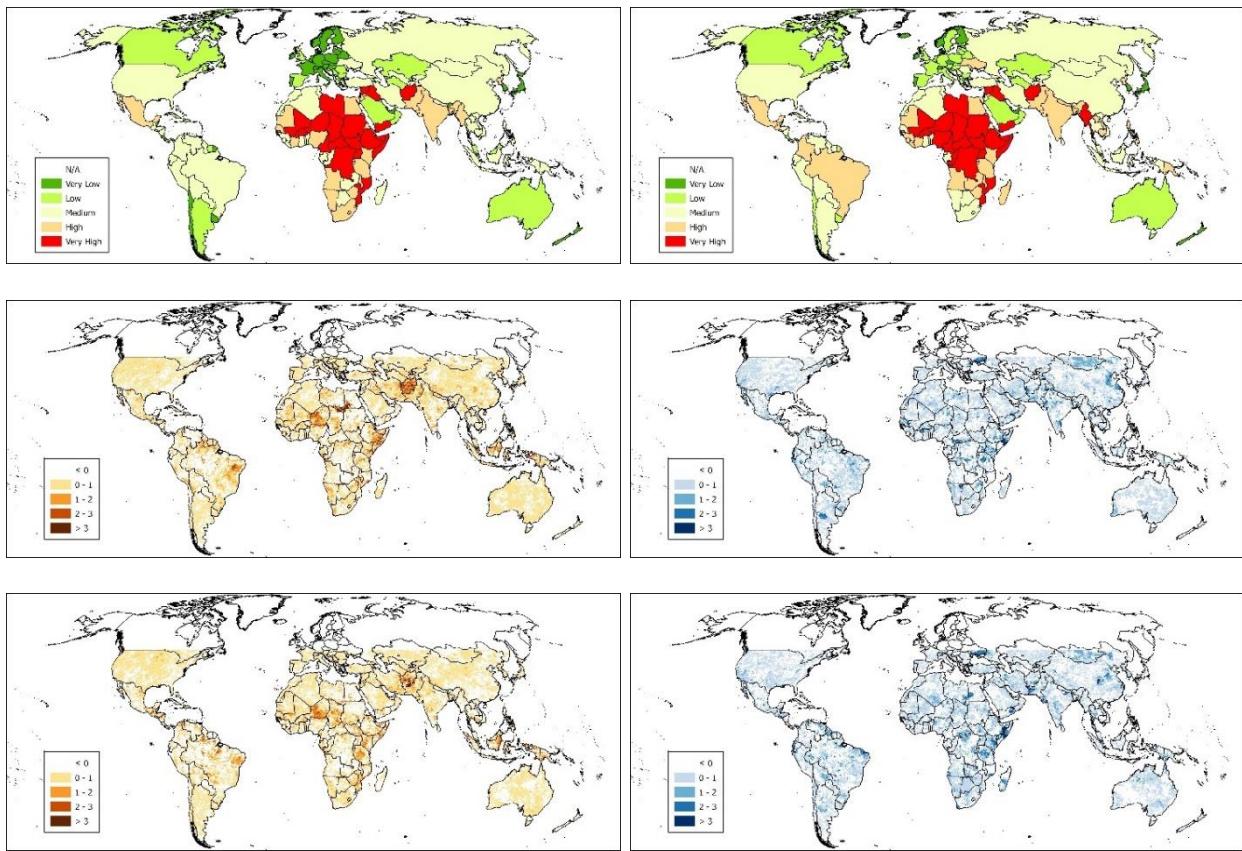
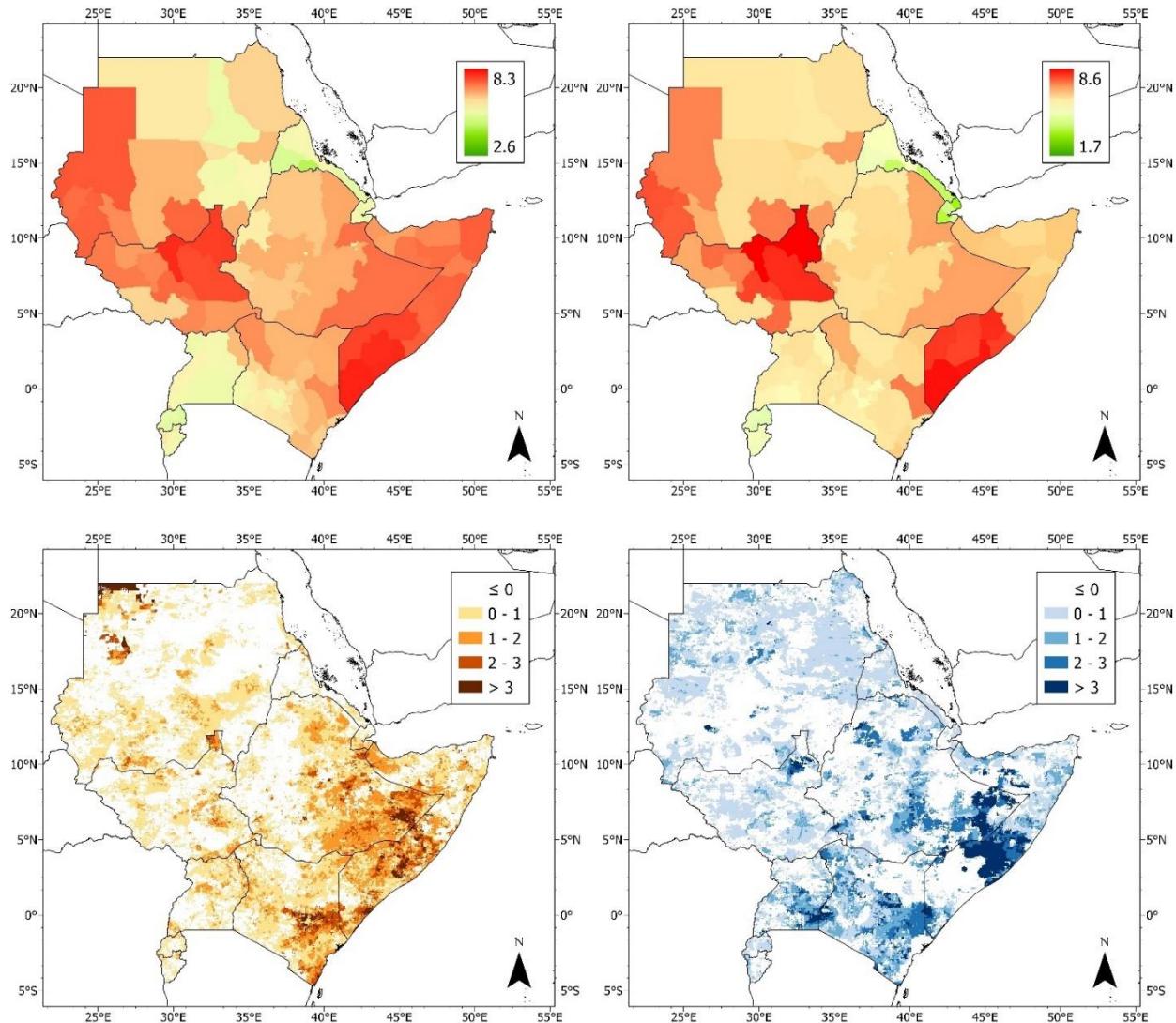


Figure 6-4. Country-level risk score (INFORM, 2020; top), suitability score for peak month (middle) and peak season (bottom), for droughts (left) and floods (right)

Considerable variance in risk in the Greater Horn of Africa is evident at the region/province level, with southern Somalia, northeastern South Sudan, and Darfur exhibiting high risk and much of Burundi, Djibouti, Eritrea, and Rwanda exhibiting relatively low risk (Figure 6-5). Suitability scores for droughts are particularly strong in eastern Ethiopia, Somalia, and southeastern Kenya at the monthly scale, although scores at the seasonal scale are notably lower in these regions, suggesting that potential anticipatory action programs in these areas may need to consider short-term water supply shocks as well as longer-term seasonal drought. These same regions demonstrate strong suitability scores for floods, although the dominant timescale switches, with seasonal suitability exceeding monthly suitability. Although eastern Ethiopia and Somalia are generally arid, they are still prone to riverine flooding, including flash floods driven by local

precipitation (OCHA, 2020), making them prime locations for the implementation of flood-based anticipatory action programs.



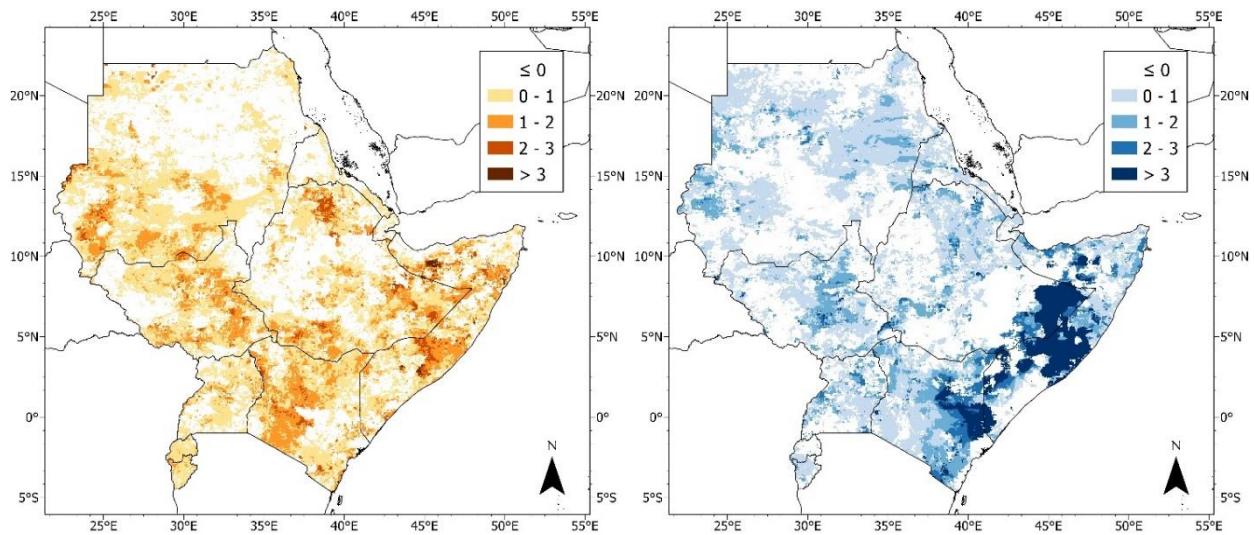
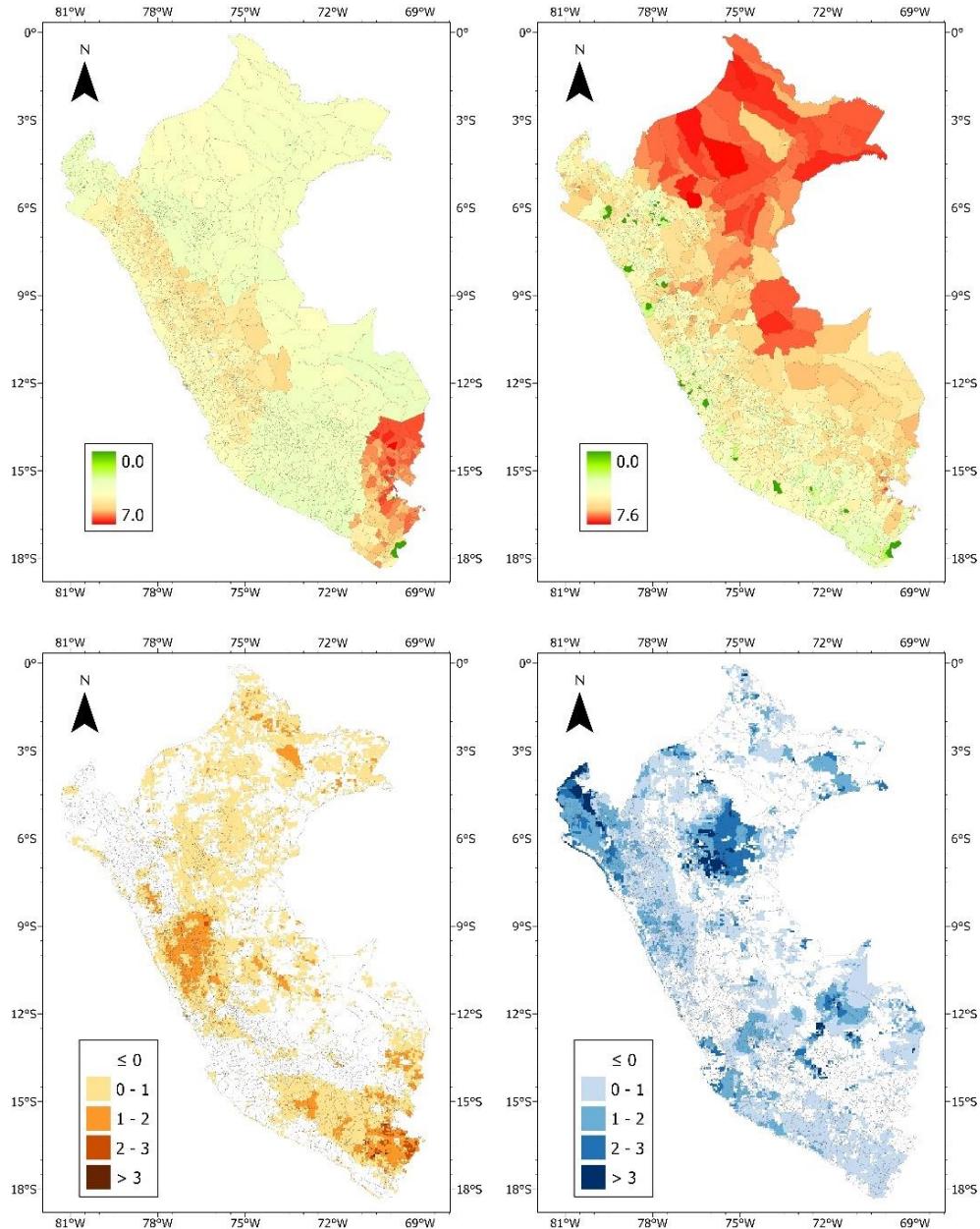


Figure 6-5. Greater Horn of Africa: risk score (INFORM, 2020; top), suitability score for peak month (middle) and peak season (bottom), for droughts (left) and floods (right)

Finally, in Peru, risk is generally much higher for floods than for droughts, except in the far southeastern part of the country (Figure 6-6). Flood risk is primarily focused on the Amazon basin in the north and east, although there exists some risk of coastal flooding in the northwest as well. Owing to the high levels of predictability in the northwest—primarily due to the effects of the El Niño-Southern Oscillation—suitability scores are generally higher than in the Amazon basin, where forecast skill is relatively poor (see Figure 6-3). Indeed, the Red Cross already has an EAP for extreme rainfall in northwestern Peru (Bazo et al., 2019; Lala et al., 2021b). Regarding drought, suitability is focused primarily in the central Andes and the extreme southeast, owing mainly to high forecast skill in the former and high risk scores in the latter.



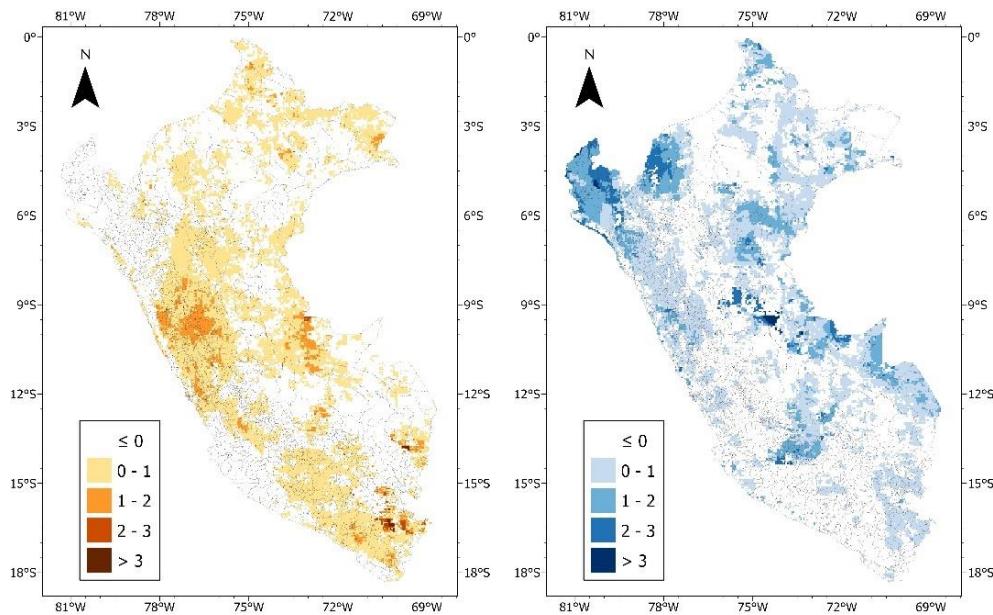


Figure 6-6. Peru: risk score (top), suitability score for peak month (middle) and peak season (bottom), for droughts (left) and floods (right)

6.4 Discussion

This study serves to address a major gap in global disaster risk reduction, specifically the inclusion of anticipation in risk analysis. The primary source of this paper’s risk data—the INFORM project—has recently called for the integration of early warning systems with hazard analysis in order to more effectively identify and mitigate future disaster risk (OCHA, 2021). Incorporating open-source and global forecasts into a globally standardized risk metric can improve the geographic determination and sustainability of future anticipatory action programs.

In general, the primary driver for suitability is forecast skill, with risk scores playing mainly an amplifying or attenuating role in the overall suitability scores. While risk scores are generally comparable between droughts and floods—owing to identical vulnerability and lack of coping capacity sub-scores—forecast skill and suitability scores are more generally distinct by disaster type. Indeed, the only major region in which suitability scores are high for both droughts and floods

is East Africa, and the precise locations within this region still vary by timescale (monthly or seasonally). Given that anticipatory action programs are generally tailored to a single hazard type, however, there exists a wide range of locations that may be suitable for these programs. For drought, much of the Sahel, Afghanistan, Indonesia, and eastern Brazil demonstrate high suitability scores, while for floods, coastal Ecuador and Peru, southern Ukraine, and some western regions of South Asia are suitable. Given that this study only considers two global forecasts, further studies may benefit from regional or local forecasts, or the inclusion of other forecast models. It should be noted, however, that the dominance of forecast skill is partially the result of how the suitability score is defined (i.e., equal weighting of forecast skill and risk score); other formulations may yield different results.

The subnational analyses of the Greater Horn of Africa and of Peru also provide insights for disaster management. In the Greater Horn of Africa, suitability scores for droughts are generally higher at the monthly scale than at the seasonal scale, suggesting that projects may be suited for relatively short-term supply shocks. Despite the frequent focus on total seasonal precipitation in studying the agro-economic impacts of water shortages in East Africa, more recent studies have found that the start of the rainy season—a subseasonal process—has a strong influence on crop yields (Lala et al., 2021a) and market price at harvest (Davenport et al., 2021), and that even short-term water shortages can lead to conflict (Maystadt and Ecker, 2014). Regarding flooding in Peru, we reiterate the presence of the current EAP for coastal extreme rainfall (Bazo et al., 2019), while also suggesting options for improved forecasts in the high risk yet low predictability parts of the Amazon basin (Keating et al., 2021).

Despite the use of case studies with subnational data to demonstrate more localized and tailored suitability, there are still data limits to conducting a global analysis. Although district-level data is available in a data-rich country like Peru, our case study in the Greater Horn of Africa is limited to the regional or provincial level which, although suitable for widespread disasters like droughts, may be too granular for in-depth flood analysis (Ward et al., 2015). Nevertheless, high resolution, skilful forecasts may allow for proper geographic targeting.

Finally, available historical forecast records (1982-2020 for NCEP CFSv2 and 1993-2020 for ECMWF SEAS5) also limit the applicability of extreme event analysis. For example, considering the 95th percentile of precipitation yields only ~2 instances of exceedance over the entire study period; the probability distributions inferred may thus misrepresent the extremes owing to this small sample size. This is especially apparent when mapping skill and suitability scores in the Greater Horn of Africa (Figure 6-5); the stark gradient in suitability in parts of Ethiopia and Somalia is likely due to a discrete change in the peak month from April to October with a corresponding sudden increase in forecast skill. Such stark gradients, especially in dry areas and for extreme events, must therefore be taken with a degree of skepticism. The use of multiple forecasts serves to partially mitigate this; however, insufficient forecast skill due to small sample sizes remains, and most regions of the world demonstrate relatively poor skill in our analysis. Climate change is also a factor; although the forecasts are bias corrected based on historic data, long-term changes in the climate may substantially change the distribution of precipitation in many regions (Cannon et al., 2015). The results of this work should therefore be limited to the short term and not be inferred in long-term climate projections.

6.5 Conclusion

This study advances the field of anticipatory action for humanitarian aid by integrating predictability of flood and drought hazards with a standardized risk metric to highlight geographic regions most suitable for program implementation. Results indicate that, even in the presence of limited forecast skill, some regions of the world demonstrate high levels of suitability, particularly in East Africa. Although data on risk is limited in spatial resolution, we find that forecast skill is a strong driver of suitability, suggesting that improvements in forecasts—via increases in skill and improvements in spatial resolution—may enhance the proper geographic targeting despite limits in risk data. On the other hand, higher resolution risk data—especially at the community scale—could be equally or more informative than forecast skill in certain contexts. Future studies may thus aim to improve forecasts, consider climate change, or investigate new regions with subnational data on risk, such that the emerging field of anticipatory action can become increasingly valuable for at-risk communities.

Chapter 7. Summary and recommendations

The steady increase in the costs—both social and economic—of climate-related disasters is a significant obstacle to global economic development, particularly in low- and middle-income countries such as Ethiopia and Peru where high levels of exposure and vulnerability leave many stuck in poverty traps. This dissertation demonstrates how hydroclimatic prediction, particularly for extreme events such as floods, droughts, and delayed onsets of rain, can act as risk reduction mechanisms for farmers and disaster relief agencies worldwide. Although this research is predominantly focused on Ethiopia and Peru, broader insights into the role of hydroclimatic prediction for disaster risk reduction may be applied to other locations, particularly those in low- or middle-income countries.

This dissertation first considers the development of an empirical, regression-based forecast of rainy season onset for Ethiopia, demonstrating skill comparable to that of dynamic models, but at a finer spatial resolution (Chapter 2). These forecasts have an ability to avoid false onsets that put farmers at risk, even though predictive skill is only moderate (i.e., reduction in mean error of 20-40% relative to climatology). Most importantly, these forecasts have been developed side-by-side with farmers, avoiding ambiguities related to onset itself—which is difficult to define—and with communication of probabilistic information. Communications with partners in the field have confirmed that these forecasts were well-received by farmers and agricultural extension agents. Indeed, onset forecasts were not originally included as part of the project, but rather arose from communications with farmers who suggested that such predictions could help them cope with an increasingly erratic start of season (for more information on the communication aspect of these forecasts, we refer readers to Alexander, 2021).

Although the onset forecast developed in Chapter 2 demonstrates skill and an ability to avoid false onset, there are no measurements of direct effects on crop yield or wider economic outcomes resulting from the adoption of the forecast. This research thus extends the forecast by using it to test alternative planting times for maize compared to a baseline scenario typically used by farmers (Chapter 3). Results indicate that forecast-informed planting has strong potential for yield gains in drier parts of Ethiopia, where precipitation is variable and onset erratic. Moreover, reductions in interannual yield variance are found countrywide, suggesting the forecast can be beneficial even in areas where rainfall is reliable, and likely more so if onset definitions are properly tailored to a wetter climatology. Given that farmers in general are risk-averse, the prospects of not just increasing mean yields but reducing variance as well results in especially positive gains when evaluating risk-averse performance metrics, again indicating that there is value in onset forecasts.

Despite evidence that onset forecasts can increase maize yields and decrease interannual variance, the wider economic effects of this practice are still not well understood. Chapter 4 therefore extends the yield gains associated with Chapter 3 by incorporating them into an economy-wide multimarket model of Ethiopia. Compared to a baseline scenario, the forecast-informed yield gains scenario results in lower maize prices, lower poverty rates, higher calorie consumption, and higher GDP overall. More generally, by disaggregating the economic model from an annual timestep to a seasonal one, broader insights into the seasonality of economic activity in the country are achieved. This is particularly relevant for the spring *Belg* season, which is often disregarded or lumped into annually aggregated analyses. The more volatile rains of the *Belg* result in lower yields, higher prices, and more interannual variance in a variety of economic indicators, highlighting the importance of the season in future modeling and relief efforts.

For a contrasting institutional and socioeconomic situation, a forecast-based financing and early action framework for floods and droughts was investigated, using a case study in Peru, a middle-income country. A sensitivity analysis of an operational early action protocol (EAP), developed by the Red Cross, for extreme rainfall in coastal Peru is conducted for a variety of forecasts, spatiotemporal resolutions, economic indicators, and trigger thresholds (Chapter 5). Broadly speaking, this research highlights the importance of the ratio of avoidable losses to the cost of the project, finding it more influential in optimizing early actions than the skill of the forecasts or the decision maker's level of risk aversion. Moreover, the study finds considerable value in seasonal and subseasonal forecasts for preparation, to the extent that the financial returns of action depend more on the willingness to act at these scales than willingness to act at the short-term, weekly scale. We therefore recommend that future EAPs incorporate these longer lead forecasts, as well as continue monitoring and evaluation of current projects to increase knowledge regarding the expected benefits of a project relative to its costs.

Having defined and evaluated a role for seasonal and subseasonal forecasts in early action, this dissertation concludes by coupling this information with risk metrics—defined with socioeconomic data—to identify the suitability of geographic locations for anticipatory actions (Chapter 6). Despite the highly variable spatial resolution of available risk data, ranging from country to district (level 2) scales, suitability primarily follows forecast skill, given its high spatial variability. General analysis can thus be conducted at the resolution of the forecasts, which can be downscaled to local data as needed. Applying this analysis to the case studies of Ethiopia and Peru, we find particular suitability for flood early action in northwestern Peru (confirming the importance of our work in Chapter 5), both drought and flood early action in eastern Ethiopia, and drought early action in northern Ethiopia (confirming the importance of our work in Chapters 2-

4). We caution, however, that cost is not incorporated into this evaluation, and that the spatial resolution of risk data, while not as important as that of forecast skill, is still a limiting factor in targeting specific communities.

Finally, returning to the three questions outlined in the introduction to this dissertation, we consider how the research described here fits into a general understanding of the role of forecasts and climate information as valuable risk reduction mechanisms for end-users:

7.1 How can forecasts and climate information at sub-seasonal to seasonal scales be tailored to provide value to end users?

Five different modeling techniques are investigated in this dissertation: statistical regressions (Chapter 2), crop models (Chapter 3), economic models (Chapter 4), sensitivity analyses of multi-stage action protocols (Chapter 5), and combined forecast and socioeconomic modeling (Chapter 6). In each case, partnerships with local specialists and end users are leveraged, ensuring a clear definition of user needs and a robust transfer of knowledge. Differences in end users are also considered by conducting research in two separate case studies. In the Ethiopia case study, subseasonal onset forecasts are provided to farmers to potentially avoid the risks of planting sub-optimally. This investigation of planting times is then extended by considering impacts on maize yields and additionally applying a partial equilibrium model to quantify the overall potential economic effects of these forecasts on the country as a whole. In Peru, this work explicitly considers actual financial returns from the use of extreme rainfall forecasts at multiple spatial and temporal scales—including at the seasonal and subseasonal level—and subsequently incorporates socioeconomic conditions in consideration of risk and forecast-informed mitigation suitability. Overall, the use of multiple modeling techniques, all in coordination and co-evaluation with end

users, presents a broad array of potential uses for seasonal and subseasonal hydroclimatic information.

7.2 What are the strengths and weaknesses of various forecast valuation techniques?

Over the five different modeling techniques used in this dissertation, benefits are explored via qualitative, quantitative, and economic valuation, explicitly addressing the need for each as outlined in Soares et al. (2018). The forecast skill demonstrated in Chapter 2, lacking quantifiable value in terms of crop yields or economic effects, is extended to quantify yield gains (Chapter 3) and broader economic insights (Chapter 4), while it also indicates qualitative value via its positive reception from farmers. The globally available forecasts applied in Chapters 5 and 6 are evaluated from a direct financial perspective (Chapter 5) and from a more qualitative perspective by considering their overlap with broader flood and drought risk categorization (Chapter 6). Forecast application and performance are therefore examined from a variety of perspectives, with subsequent chapters building upon gaps in the previous chapters. Future studies may extend this work by explicitly capturing the behavior of beneficiaries, including but not limited to the effect of false alarms on forecast effectiveness, the role of forecast uptake and social interaction (e.g., via agent-based modeling), and decision theory-based approaches to understanding the role of forecasts on individual decision making.

7.3 How do different institutions and socioeconomic conditions affect the use and value of forecasts in disaster contexts?

The work covered in this dissertation includes projects that were developed and executed in collaboration with agricultural extension agents, meteorological agencies, research institutions, and relief agencies, in addition to other universities. Case studies explicitly consider

socioeconomic conditions by including a low-income agriculture country (Ethiopia) and a middle-income urbanized one (Peru). The work in Ethiopia is developed in co-evaluation with farmers (Chapter 2), agricultural modelers and extension agents (Chapter 3), and economic research institutions (Chapter 4), all within the context of a seasonally driven agricultural society. Meanwhile, the work in Peru is conducted in collaboration with the Red Cross and the Peruvian meteorological and hydrological society, focusing specifically on efforts to optimize early action protocols arising from challenges identified in a 2015-16 pilot project (Chapter 5). Such forecast-based early action considerations are then expanded to explicitly include socioeconomic factors of risk at the country level globally and at the subnational level for Ethiopia and Peru (Chapter 6). In short, these considerations contribute to the overall understanding of forecasts by adding to a corpus of work that often focuses disproportionately on wealthy countries. Future studies should continue to investigate such differences in institutional and socioeconomic contexts as part of a holistic approach to disaster risk reduction.

7.4 Final words

Seasonal and subseasonal hydroclimatic prediction is a critical component of disaster management with promising prospects for risk reduction, as demonstrated by the range of forecasts, actions, institutional and socioeconomic contexts, spatiotemporal factors, and valuation techniques explored in this dissertation. By connecting short-term weather phenomena with long-term climate projections, seasonal and subseasonal forecasts act as an effective—yet often underutilized—bridge in timescales that can be of substantive use to farmers, relief agencies, and others invested in disaster risk reduction. Future work should continue to explore the utility of seasonal and subseasonal hydroclimatic prediction across sectors, particularly in the context of climate change which will present new challenges in an ever-developing world.

References

- Abdi, H. (2003). "Partial Least Squares (PLS) Regression." In: Lewis-Beck, M., Bryman, A., and Futing, T. (Eds.). Encyclopedia of Social Sciences Research Methods, Thousand Oaks (CA): Sage
- Ahmadalipour, A., and Moradkhani, H. (2018). Multi-dimensional assessment of drought vulnerability in Africa: 1960-2100. *Science of the Total Environment*, 644, 520-535, doi:10.1016/j.scitotenv.2018.07.023
- Alexander, S. O. (2021). Evaluation, communication, and integration of climate information across scales to foster local decision-making and support community resilience. Doctoral dissertation, University of Wisconsin-Madison
- Alexander, S., Wu, S., and Block, P. (2019). Model selection based on sectoral application scale for increased value of hydroclimate prediction information. *J. Water Res. Planning and Mgmt.*, 145, doi:10.1061/(ASCE)WR.1943-5452.0001044
- Alexander, S., and Block, P. (2021). Exploring the integration of climate forecast information into local-level decision making using an agent-based model. In review.
- Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). Crop evapotranspiration – Guidelines for computing crop water requirements – FAO irrigation and drainage paper 56 FAO. Rome 300:D05109
- Araya, A., Stroosnijder, L., Habtu, S., Keesstra, S. D., Berhe, M., and Hadgu, K. M. (2012). Risk assessment by sowing date for barley (*Hordeum vulgare*) in northern Ethiopia. *Agricultural and Forest Meteorology*, 154-155, 30-37, doi:10.1016/j.agrformet.2011.11.001

Arrow, K. J. (1965). "Aspects of the Theory of Risk Bearing". *The Theory of Risk Aversion*. Helsinki: Yrjo Jahnssonin Saatio. Reprinted in: *Essays in the Theory of Risk Bearing*, Markham Publ. Co., Chicago, 1971, 90–109.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1998). Coherent measures of risk. *Mathematical Finance*, 9, 203-228, doi:10.1111/1467-9965.00068

Aybar, C. (2018). Flood SENAMHI GitHub repository,
<https://github.com/csaybar/slides/tree/master/flood>

Aybar, C., Fernandez, C., Huerta, A., Lavado, W., Vega, F., and Felipe-Obando, O. (2019). Construction of a high-resolution gridded rainfall dataset for Peru from 1981 to the present day. *Hydrological Sci. J.*, 65, doi:10.1080/02626667.2019.1649411

Ayele, S. (2021). The resurgence of agricultural mechanisation in Ethiopia: rhetoric or real commitment? *J. Peasant Studies*, p. 1-21, doi:10.1080/03066150.2020.1847091

Bachewe, F., Minten, B., Taffesse, A. S., Pauw, K., Cameron, A., and Endaylalu, T. G. (2018). Farmers' grain storage and losses in Ethiopia. IFPRI Working Paper 115, March 2018

Bakker, C., Zaitchik, B. F., Siddiqui, S., Hobbs, B. F., Broaddus, E., Neff, R. A., Haskett, J., and Parker, C. L. (2018). Shocks, seasonality, and disaggregation: modeling food security through the integration of agricultural, transportation, and economic systems. *Agricultural Systems*, 164, 164-184, doi:10.1016/j.agsy.2018.04.005

Barnes, L. R., Benight., C. C., Gruntfest, E. C., Hayden, M. H., Schultz, D. M., and Benight, C. (2007). False alarms and close calls: a conceptual model of warning accuracy. *Weather Forecasting*, 22, 1140-1147, doi:10.1175/WAF1031.1

Bar-Shira, Z., Just, R. E., and Zilberman, D. (1997). Estimation of farmers' risk attitude: an econometric approach. *Agricultural Econ.*, 17, 211-222, doi:10.1111/j.1574-0862.1997.tb00475.x

Bazo, J., Lorenzo, M.d.l.N, and Porfirio de Rocha, R. (2013). Relationship between monthly rainfall in NW Peru and tropical sea surface temperature. *Advances in Meteorology*, 2013, doi:10.1155/2013/152875

Bazo, J., Singh, R., Destrooper, M., and Coughlan de Perez, E. (2019). Chapter 18 – Pilot experiences in using seamless forecasts for early action: the “Ready-Set-Go!” approach in the Red Cross. Chapter in Sub-Seasonal to Seasonal Prediction, Ed. A. W. Robertson and F. Vitart, pp. 387-398, doi:10.1016/B978-0-12-811714-9.00018-8

Beck, H. E., Van Dijk, A.I.J.M., Levizzani, V., Schellekens, J., Miralles, D.G., Martens, B.,and De Roo, A. (2017a). MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci.*, 21, 589-615, doi:10.5194/hess-21-589-2017

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., Brocca, L., Pappenberger, F., Huffman, G. J., and Wood, E. F. (2017b). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modelling. *Hydrol. Earth Syst. Sci.*, 21, 6201–6217, doi:10.5194/hess-21-6201-2017, 2017

Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., Van Dijk, A. I. J. M., Huffman, G. J., Adler, R. F., and Wood. E. F. (2019). Daily evlatuion of 26 precipitatio datasets using Stage-IV gauge-radar data for the CONUS. *Hydrol. Earth Syst. Sci.*, 23, 207-244, doi:10.5194/hess-23-207-2019

Berhane, F., and Zaitchik, B. (2014). Modulation of daily precipitation over East Africa by the Madden-Julian Oscillation. *Journal of Climate*, 27, 6016-6034, doi:10.1175/JCLI-D-13-00693.1

Berhane, G., Dereje, M., Minten, B., Tamru, S. (2017). The rapid – but from a low base – uptake of agricultural mechanization in Ethiopia: Patters, implications and challenges. International Food Policy Research Institute, ESSP Working Paper 105, April 2017

Bischiniotis, K., van den Hurk, B., Coughlan de Perez, E., Veldkamp, T., Guimarães Nobre, G., and Aerts, J. (2019). Assessing time, cost and quality trade-offs in forecast-based action for floods. *Int. J. Disaster Risk Reduction*, 40, doi:10.1016/j.ijdrr.2019.101252

Bischiniotis, K., de Moel, H., van den Homberg, M., Couasnon, A., Aerts, J., Guimarães Nobre, G., Zsoter, E., and van den Hurk, B. (2020). A framework for comparing permanent and forecast-based flood risk-reduction strategies. *Sci. Total Environment*, 720, doi:10.1016/j.scitotenv.2020.137572

Block, P. J. (2006). Integrated management of the Blue Nile Basin in Ethiopia: Precipitation forecast, hydropower, and irrigation modeling. Thesis: University of Colorado at Boulder.

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, doi:10.1111/j.1574-0862.2008.00322.x

Borgomeo, E., Hall, J. W., Salehin, M. (2017). Avoiding the water-poverty trap: insights from a conceptual human-water dynamical model for coastal Bangladesh. *Int. J. Water Resources Development*, 34, 900-922, doi:10.1080/07900627.2017.1331842

Box, G. E. P., and Cox, D. R. (1964). An analysis of transformations. *J. Royal Statistical Society, Series B*. 26 (2): 211–252. JSTOR 2984418. MR 0192611.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45 (1), 5–32, doi:10.1023/A:1010933404324

Brier, G. W. (1950). Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review*, 78: 1–3, doi:10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2

Butler, D. (2015). Aid burst lifts people out of extreme poverty. *Nature*, 521, 269, doi:10.1038/nature.2015.17560

C3S – Copernicus Climate Change Service (2017). ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), 20 August 2020. <https://cds.climate.copernicus.eu/cdsapp#!/home>

Cabot Venton, C. (2018). Economics of resilience to droughts: Ethiopia analysis. United States Agency for International Development, 43 pp., https://www.usaid.gov/sites/default/files/documents/1867/Ethiopia_Economics_of_Resilience_Final_Jan_4_2018_-_BRANDED.pdf.

Cabot Venton, C. (2018). Economics of Resilience to Droughts: Ethiopia Analysis. United States Agency for International Development, Washington, D.C.,

https://www.usaid.gov/sites/default/files/documents/1867/Ethiopia_Economics_of_Resilience_Final_Jan_4_2018_-_BRANDED.pdf

Cannon, A. J., Sobie, S. R., and Murdock, T. Q. (2015). Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes? *J. Climate*, 28 (17), doi:10.1175/JCLI-D-14-00754.1

CIA – Central Intelligence Agency World Factbook (2020). Ethiopia. The World Factbook, <https://www.cia.gov/the-world-factbook/countries/ethiopia/>

Coughlan de Perez, E., van den Hurk, B., van Aalst, M. K., and Jongman, B. (2015). Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts. *Nat. Hazards Earth Syst. Sci.*, 15, 895–904, doi:10.5194/nhess-15-895-2015

Craven, P., and Wahba, G. (1979). Smoothing noisy data with spline functions. *J. Numer. Math.*, 31, 377-403

CSA – Central Statistical Agency of Ethiopia (2016). Agricultural Sample Survey. *Stat. Bull.* 1, 1–111.

Darabant, A., Habermann, B., Sisay, K., Thurnher, C., Worku, Y., Damtew, S., Lindtner, M., Burrell, L., and Abiyu, A. (2020). Farmers' perceptions and matching climate records jointly explain adaptation responses in four communities around Lake Tana, Ethiopia. *Climatic Change*, 163, 481-497, doi:10.1007/s10584-020-02889-x

Davenport, F. M., Shukla, S., Turner, W., Funk, C., Krell, N., Harrison, L., Husak, G., Lee, D., and Peterson, S. (2021). Sending out an SOS: using start of rainy season indicators for market

price forecasting to support famine early warning. Environmental Research Letters, 16, doi:10.1088/1748-9326/ac15cc

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F. (2011). The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* 137, 553–597, doi:10.1002/qj.82

Di Baldassarre, G., Wanders, N., AghaKouchak, A., Kuil, L., Rangecroft, S., Veldkamp, T. I. E., Garcia, M., van Oel, P. R., Breinl, K., and Van Loon, A. F. (2018). Water shortages worsened by reservoir effects. *Nature Sustainability*, 1, 617-622, doi:10.1038/s41893-018-0159-0

Diao, X., and Pratt, A. N. (2007). Growth options and poverty reduction in Ethiopia – An economy-wide model analysis. *Food Policy*, 32, 205-228, doi:10.1016/j.foodpol.2006.05.005

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H., and Ceccato, P. (2018). Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Q. J. R. Meteorol. Soc.*, 144 (Suppl. 1), 292– 312, doi:10.1002/qj.3244

Dorosh, P. and Minten, B. (2020). Ethiopia's Agrifood System: Past Trends, Present Challenges, and Future Scenarios. Washington, DC: International Food Policy Research Institute. doi:10.2499/9780896296916

Drechsler, M., and Soer, W. (2016). Early warning, early action: the use of predictive tools in drought response through Ethiopia's productive safety net programme. World Bank Policy Research Working Paper No. 7716, June 2016, doi:10.1596/1813-9450-7716

Dunning, C. M., Black, E. C. L., and Allan, R. P. (2016). The onset and cessation of seasonal rainfall over Africa. *J. Geophys. Res. Atmos.*, 121, 11405-11424, doi:10.1002/2016JD025428

ECMWF – European Centre for Medium-Range Weather Forecasts (2021). The ECMWF Integrated Forecasting System (IFS). ecmwf.int/en/forecasts/

Ethiopia CSA – Central Statistical Agency (2015). Report on farm management practices. Addis Ababa, August 2015

Ethiopia CSA – Central Statistical Agency (2017). Report on area and production for major crops. Addis Ababa, April 2017

Ethiopia CSA – Central Statistical Agency (2018). Ethiopian household consumption – expenditure (HCE) survey. Addis Ababa, January 2018

Ewbank, R., Perez, C., Cornish, H., Worku, M., and Woldetsadik, S., (2019). Building resilience to El Niño-related drought: experiences in early warning and early action from Nicaragua and Ethiopia. *Disasters*, 43, S345-S367, doi:10.1111/dis.12340

FAO-UNESCO (1988), Soil map of the world, revised legend. *World Resources Report* 60:138

FAO – Food and Agriculture Organization of the United Nations (2010). “Crop calendar – an informational tool for seed security.”

<http://www.fao.org/agriculture/seed/cropcalendar/welcome.do>, accessed 20 July 2020

Ferguson, C. R., Pan, M., and Oki, T. (2018). The effect of global warming on future water availability: CMIP5 synthesis. Water Resources Research, 54, 7791-7819, doi:10.1029/2018WR022792

Funk, C. C., Peterson, P. J., Landsfeld, M. F., Pedreros, D. H., Verdin, J. P., Rowland, J. D., Romero, B. E., Husak, G. J., Michaelsen, J. C., and Verdin, A. P. (2014). A quasi-global precipitation time series for drought monitoring: U.S. Geological Survey Data Series 832, 4 p.
<http://pubs.usgs.gov/ds/832/>

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J.C. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci. Data, 2, doi:10.1088/1748-9326/ac15cc

Funk C., Harrison, L., Alexander, L., Peterson, P., Behrang, A., and Husak, G. (2019). Exploring trends in wet-season precipitation and drought indices in wet, humid and dry regions. Environmental Res. Letters, 14, doi:10.1088/1748-9326/ab4a6c

GFDRR Labs (2021). ThinkHazard! database: Water scarcity, Peru.
<https://thinkhazard.org/en/report/195-peru/DG>

Government of Ethiopia (2019). “Ethiopia’s Climate Resilient Green Economy – National Adaptation Plan.” Prepared for the United Nations Framework Convention on Climate Change, March 2019

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T. (2012). Technical note: downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods. *HESS*, 16, 3383-3390, doi:10.5194/hess-16-3383-2012

Guimarães Nobre, G., Davenport, F., Bischiniotis, K., Veldkamp, T., Jongman, B., Funk, C. C., Husak, G., Ward, P. J., and Aerts, J. (2018). Financing agricultural drought risk through ex-ante cash transfers. *Sci. Total Environment*, 653, 523-535, doi:10.1016/j.scitotenv.2018.10.406

Hansen, J. W., Mishra, A., Rao, K. P. C., Indeje, M., Ngugi, R. K. (2009). Potential value of GCM-based seasonal rainfall forecasts for maize management in semi-arid Kenya. *Agricultural Systems*, 101 (1–2), 80–90, doi:10.1016/j.agsy.2009.03.005

Harris, I., Osborn T. J., Jones, P., and Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, 7, 109, doi:10.1038/s41597-020-0453-3

Hastenrath, S., Polzin, D., and Mutai, C. (2011). Circulation mechanisms of Kenya rainfall anomalies. *J. Clim.*, 24(2), 404–412, doi:10.1175/2010JCLI3599.1

Hirvonen, K., Taffesse, A. S., and Worku, I. (2015). Seasonality and household diets in Ethiopia. IFPRI Working Paper 74, May 2015

Hoogenboom, G., Jones, J.W., Wilkens, P.W., Porter, C.H., Boote, K.J., Hunt, L.A., Singh, U., Lizaso, J.L., White, J.W., Uryasev, O., Royce, F.S., Ogoshi, R., Gijsman, A.J., Tsuji, G.Y., and Koo, J. (2012). Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.5 [CD-ROM]. University of Hawaii, Honolulu, Hawaii

Horton, R., and Lo, S. (2015). Planetary health: a new science for exceptional action. *The Lancet*, 386, doi:10.1016/S0140-6736(15)61038-8

Huang, L., Frijters, P., Dalziel, K., and Clarke, P. (2018). Life satisfaction, QALYs, and the monetary value of health. *Social Science & Medicine*, 211, 131-136, doi:10.1016/j.socscimed.2018.06.009

IFRC—International Federation of Red Cross and Red Crescent Societies (2020). Forecast-based financing: our projects. <https://www.forecast-based-financing.org/our-projects/>

Imfield, N., Schuler, C. B., Correa Marrou, K. M., Jacques-Coper, M., Sedlmeier, K., Gubler, S., Huerta, A., and Brönnimann, S. (2019). Summertime precipitation deficits in the southern Peruvian highlands since 1964. *Int. J. Climatology*, 39, 4497-4513, doi:10.1002/joc.6087

INDECI (2017). “Boletín Estadístico Virtual de la Gestión Reactiva” (Report). No. 07, Año 4, Jul. 2017

INEI – Instituto Nacional de Estadística e Informática (2017). Bases de datos, <https://www.inei.gob.pe/>

INFORM (2020). INFORM Risk Index 2021. <https://drmkc.jrc.ec.europa.eu/inform-index>. INFORM is a collaboration of the Inter-Agency Standing Committee Reference Group on Risk, Early Warning and Preparedness and the European Commission. The European Commission Joint Research Centre is the scientific lead of INFORM.

International Research Institute for Climate and Society (IRI), Michigan State University (MSU), and HarvestChoice (IFPRI) (2015). Global High-Resolution Soil Profile Database for Crop Modeling Applications. Harvard Dataverse, V2, doi:10.7910/DVN/1PEEY0

IPCC – The Intergovernmental Panel on Climate Change (2014). Climate Change 2014: Impacts, Adaptation, and Vulnerability. Working Group II contribution to the Fifth Assessment Report of the IPCC. New York: Cambridge University Press, © Intergovernmental Panel on Climate Change 2014

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S. P. E., Mogensen, K., Zuo, H., and Monge-Sanz, B. M. (2019). SEAS5: the new ECMWF seasonal forecast system. Geosci. Model Dev. 12, 1087-1117, doi:10.5194/gmd-12-1087-2019

Jury, M. R. (2011). Meteorological scenario of Ethiopian floods in 2006-2007. Theoretical and Applied Climatology, 104, 209-219, doi:10.1007/s00704-010-0337-0

Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. Educational and Psychological Measurement, XX(1), 141-151

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S-K., Hnilo, J. J., Fiorino, M., and Potter, G. L. (2002). NCEP-DOE AMIP-II Reanalysis (R-2). 1631-1643, Nov 2002, Bulletin of the American Meteorological Society

Keating, C., Lee, D., Bazo, J., and Block, P. (2021). Leveraging multi-model season-ahead streamflow forecasts to trigger advanced flood preparedness. Natural Hazards and Earth System Sciences, 21, 2215-2231, doi:10.5194/nhess-21-2215-2021

Kelman, Ilan. 2013. Disaster Mitigation is Cost Effective. World Bank, Washington, DC. © World Bank. <https://openknowledge.worldbank.org/handle/10986/16341> License: CC BY 3.0 IGO.

Khalil, A. F., Kwon, H-H., Lall, U., Miranda, M. J., and Skees, J. (2007). El Niño-Southern Oscillation-based index insurance for floods: Statistical risk analyses and application to Peru. *Water Resources Research*, 43, doi:10.1029/2006WR005281

Kipkorir, E. C., Raes, D., Bargerei, R. J., and Mugalavai, E. M. (2007). Evaluation of two risk assessment methods for sowing maize in Kenya. *Agricultural and Forest Meteorology*, 144, 193-199, doi:10.1016/j.agrformet.2007.02.008

Kirtman, B. P., and coauthors (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, doi:10.1175/BAMS-D-12-00050.1

Korecha, D., and Barnston, A. G. (2007). Predictability of June-September rainfall in Ethiopia. *Monthly Weather Review*, 135, 628-650, doi:10.1175/MWR3304.1

Lala, J., Tilahun, S., and Block, P. (2020). Predicting rainy season onset in the Ethiopian Highlands for agricultural planning. *J. Hydrometeorology*, 21, doi:10.1175/JHM-D-20-0058.1

Lala, J., Yang, M., Wang, G., and Block, P. (2021a). Utilizing rainy season onset predictions to enhance maize yields in Ethiopia. *Environmental Research Letters*, 16, doi:10.1088/1748-9326/abf9c9

Lala, J., Bazo, J., Anand, V., and Block, P. (2021b). Optimizing forecast-based actions for extreme rainfall events. *Climate Risk Management*, 34, doi:10.1016/j.crm.2021.100374

Lanckreit, S., Frankl, A., Adgo, E., Termonia, P., and Nyssen, J. (2014). Droughts related to quasi-global oscillations: a diagnostic teleconnection analysis in North Ethiopia. International J. of Climatology, 35, 1534-1542, doi:10.1002/joc.4074

Lee, D., Ward, P., and Block, P. (2018). Attribution of large-scale climate patterns to seasonal peak-flow and prospects for predictions globally. Water Resources Research, 54, 916-938, doi:10.1002/2017WR021205

Lee, D., Ahmadul, H., Patz, J., and Block, P. (2021). Predicting social and health vulnerability to floods in Bangladesh. Natural Hazards and Earth System Sciences, 21, 1807-1823, doi:10.5194/nhess-21-1807-2021

Liben, M. F., Wortmann, S. C., and Fantaye, T. K. (2015). Risks associated with dry soil planting time in Ethiopia. International Journal of Agronomy and Agricultural Research, 6, 14-23

Liebmann, B., and Marengo, J. A. (2001). Interannual variability of the rainy season and rainfall in the Brazilian Amazon Basin. J. Clim., 14, 4308-4318, doi:10.1175/1520-0442(2001)014<4308:IVOTRS>2.0.CO;2

Lins, D. A., Gabriel, S. C., and Sonka, S. T. (1981). An analysis of the risk aversion of farm operators: an asset portfolio approach. Western J. Agricultural Economics, 6, 15-30, doi:10.22004/ag.econ.32078

Lopez, A., Coughlan de Perez, E., Bazo, J., Suarez, P., van den Hurk, B., and van Aalst, M. (2017). Bridging forecast verification and humanitarian decisions: A valuation approach for setting up action-oriented early warnings. Weather and Climate Extremes, 27, doi:10.1016/j.wace.2018.03.006

Lorenz, C., Portele, T. C., Laux, P., and Kunstmann, H. (2021). Bias-corrected and spatially disaggregated seasonal forecasts: a long-term reference forecast product for the water sector in semi-arid regions. *Earth Syst. Sci. Data*, 13, 2701–2722, doi:10.5194/essd-13-2701-2021

MacLeod, D. (2018). Seasonal predictability of onset and cessation of the east African rains. *Weather and Climate Extremes*, 21, 27-35, doi:10.1016/j.wace.2018.05.003

Marin-Ferrer, M., Vernaccini, L. and Poljansek, K. (2017). Index for Risk Management INFORM Concept and Methodology Report — Version 2017, EUR 28655 EN, doi:10.2760/094023

Matte, S., Boucher, M. A., Boucher, V., and Filion, T. C. F. (2017). Moving beyond the cost-loss ratio: economic assessment of streamflow forecasts for a risk-averse decision maker. *HESS*, 21, 2967-2986, doi:10.5194/hess-21-2967-2017

Maystadt, J. F., and Ecker, O. (2014). Extreme weather and civil war: does drought fuel conflict in Somalia through livestock price shocks? *American J. Ag. Econ.*, 96, 1157-1182, doi:10.1093/ajae/aau010

Mengistu Tsidiu, G. (2012). High-resolution monthly rainfall database for Ethiopia: Homogenization, reconstruction, and gridding. *J. of Climate*, 25, 8422-8443, doi:10.1175/JCLI-D-12-00027.1

Moron, V., Camberlin, P., and Robertson, A. W. (2013). Extracting subseasonal scenarios: An alternative method to analyze seasonal predictability of regional-scale tropical rainfall. *J. Clim.*, 26(8), 2580–2600, doi:10.1175/JCLI-D-12-00357.1

NCAR – National Center for Atmospheric Research (2019). Weather Research and Forecasting (WRF) Model. Version 4.1, github.com/wrf-model/WRF/releases/tag/v4.1

NCEP – National Centers for Environmental Prediction (2006). The North American Mesoscale (NAM) forecast system. National Oceanic and Atmospheric Administration, nco.ncep.noaa.gov/pmb/products/nam/

NDMC – National Drought Mitigation Center (2021). U.S. Drought Monitor, University of Nebraska-Lincoln,

<https://droughtmonitor.unl.edu/About/AbouttheData/DroughtClassification.aspx>

Netherlands Red Cross (2021). 510 Dashboard, <https://dashboard.510.global>

Ng, K. S. (2013). A simple explanation of partial least squares. Tech Report, Australian National University, <http://users.cecs.anu.edu.au/~kee/pls.pdf>

Nicholson, S. E. (2017). Climate and climatic variability of rainfall over eastern Africa: Climate Over Eastern Africa. *Reviews of Geophysics*, 55(3), 590–635, doi:10.1002/2016RG000544

Nicholson, S. E. (2018). The ITCZ and the seasonal cycle over equatorial Africa. *Bulletin of the American Meteorological Society*, 99, 337-348, doi:10.1175/BAMS-D-16-0287.1

OCHA – United Nations Office for the Coordination of Humanitarian Affairs (2020). Somalia: Flash floods – Oct 2020. OCHA Updates, <https://reliefweb.int/disaster/ff-2020-000221-som>

OCHA – United Nations Office for the Coordination of Humanitarian Affairs (2021). Proposal for INFORM Warning, <https://drmkc.jrc.ec.europa.eu/inform-index/INFORM-Warning>, June 2021

Peruvian Red Cross (2016). Lessons learnt and opportunities: El Niño Peru 2015-2016. Report on Forecast based Financing (FbF).

Peruvian Red Cross (2018). Plan de Acción Temprana – Fenómeno El Niño (Lluvias Extremeas). November 2018.

Philip, S., Kew, S. F., van Oldenborgh, G. J., Otto, F., O'Keefe, S., Haustein, K., King, A., Zegeye, A., Eshetu, Z., Singh, K. H. R., Jjemba, E., Funk, C., and Cullen, H. (2018). Attribution analysis of the Ethiopian Drought of 2015. *Journal of Climate*, 31, 2465-2486, doi:10.1175/JCLI-D-17-0274.1

Pratt, J. W. (1964). Risk Aversion in the small and in the large. *Econometrica*, 32 (1–2), 122–136, doi:10.2307/1913738

Rau, P., Bourrel, L., Labat, D., Melo, P., Dewitte, B., Frappart, F., Lavado, W., and Felipe, O. (2016). Regionalization of rainfall over the Peruvian Pacific slope and coast. *Int. J. Climatology*, 37, 143-158, doi:10.1002/joc.4693

Reguero, B. G., Losada, I. J., Díaz-Simal, P., Méndez, F. J., and Beck, M. W. (2015). Effects of climate change on exposure to coastal flooding in Latin America and the Caribbean. *PLOS One*, 10(7): e0133409, doi:10.1371/journal.pone.0133409

Rosell, S. (2011). Regional perspective on rainfall change and variability in the central highlands of Ethiopia, 1978–2007. *Applied Geography*, 31, 329-338, doi:10.1016/j.apgeog.2010.07.005

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Peña Mendez, M., van den Dool, H., Zhang, Q., Wang,

W., Chen, M., and Becker, E. (2014). The NCEP Climate Forecast System Version 2. *J. Climate*, doi:10.1175/JCLI-D-12-00823.1

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

Segele, Z. T., Richman, M. B., Leslie, L. M., and Lamb, P. J. (2015). Seasonal-to-interannual variability of Ethiopia/Horn of Africa monsoon. Part II: Statistical multimodel ensemble rainfall predictions. *J. Climate*, 28, 3511–3536, doi:10.1175/JCLI-D-14-00476.1.

Skaggs, R., Janetos, T. C., Hubbard, K. A., and Rice, J. S. (2012). Climate and energy-water-;and system interactions. Technical report to the U.S. Department of Energy in support of the National Climate Assessment. Pacific Northwest National Laboratory, March 2012

Skees, J. (2011). Innovations in index insurance for the poor in low-income countries. Chapter in Financial Inclusion, Innovation, and Investments: Biotechnology and Capital Markets Working for the Poor by R. D. Christy and V. L. Bogan. Singapore: World Scientific Publishing Co. Pte. Ltd.

Soares, M. B., Daly, M., and Dessai, S. (2018). Assessing the value of seasonal climate forecasts for decision-making. *WIREs Climate Change*, 9, doi:10.1002/wcc.523

Steduto, P., Hsiao, T. C., Raes, D., and Fereres, E. (2009). AquaCrop: The FAO Crop Model to Simulate Yield Response to Water: I. Concepts and Underlying Principles. *Agronomy Journal*, 101: 426–437

Tadross, M., Suarez, P., Lotsch, A., Hachigonta, S., Mdoka, M., Uganai, L., Lucio, F., Kamdonyo, D., and Muchinda, M. (2009). Growing-season rainfall and scenarios of future change in southeast Africa: implications for cultivating maize. *Climate Research*, 40: 147-161, doi:10.3354/cr00821

Taffesse, A. S., Dorosh, P., and Asrat, S. (2012). Crop production in Ethiopia: regional patterns and trends. *Ethiopia Strategy Support Program II Working Paper 16*, International Food Policy Research Institute and Ethiopia Development Research Institute.

Temam, D., Uddameri, V., Mohammadi, G., Hernandez, E. A., and Ekwaro-Osire, S. (2019). Long-term drought trends in Ethiopia with implications for dryland agriculture. *Water*, 11, 2571, doi:10.3390/w11122571

Teshome, A., and Zhang, J. (2019). Increase of extreme drought over Ethiopia under climate warming. *Advances in Meteorology*, 2019, doi:10.1155/2019/5235429

Trainor, J. E., Nagelle, D., Philips, B., and Scott, B. (2015). Tornadoes, social science, and the false alarm effect. *Weather, Climate, and Society*. 7, 333-352, doi:10.1175/WCAS-D-14-00052.1

United Nations (2015). Sendai Framework for Disaster Risk Reduction 2015-2030. UN World Conference on Disaster Risk Reduction, March 2015, Sendai, Japan. Geneva: United Nations Office for Disaster Risk Reduction

Vavrus, S. J., Wang, F., and Block P. (2021). Rainy season precipitation forecasts in coastal Peru from the North American Multi-Model Ensemble (NMME). *Int. J. Climatology*, in review.

Veldkamp, T. I. E., Wada, Y., Aerts, J. C. J. H., and Ward, P. (2016). Towards a global water scarcity risk assessment framework: incorporation of probability distributions and hydro-climatic variability. *Environmental Research Letters*, 11, doi:10.1088/1748-9326/11/2/024006

Viste, E. and Sorteberg, A. (2013). Moisture transport into the Ethiopian highlands. *Int. J. Climatol.*, 33: 249-263. doi:10.1002/joc.3409

Vitart, F., and Robertson, A. W. (2018). The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *npj Climate and Atmospheric Science*, 1, doi:10.1038/s41612-018-0013-0

von Neumann, J., and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press

Wambui, C. (2019). “Kenyan farmers trust tradition over tech to predict the weather.” Thomas Reuters Foundation, 11 February 2019, <https://news.trust.org/item/20190211065115-ywdmq/?source=spotlight>

Ward, P. J., Blauthut, V., Bloemendaal, N., Daniell, E. J., De Ruiter, C. M., Duncan, J. M., Emberson, R., Jenkins, F. S., Kirschbaum, D., Kunz, M., Mohr, S., Muis, S., Riddell, A. G., Schäfer, A., Stanley, T., Veldkamp, I. E., and Hessel, W. C. (2020). Review article: Natural hazard risk assessments at the global scale, *Nat. Hazards Earth Syst. Sci.*, 20, 1069–1096, doi:10.5194/nhess-20-1069-2020

Ward, P. J., Jongman, B., Salamon, P., Simpson, A., Bates, P., de Goeve, T., Muis, S., Coughlan de Perez, E., Rudari, R., Trigg, M. A., and Winsemius, H. C. (2015). Usefulness and limitations of global flood risk models. *Nature Climate Change*, 5, 712-715, doi:10.1038/nclimate2742

Wilks, D. S. (1995). Statistical Methods in the Atmospheric Sciences, Elsevier, New York

WMO (2018). WWRP/WCRP Sub-seasonal to Seasonal Prediction Project (S2S) Phase II Proposal (November 2018 – December 2023). WWRP 2018-4, WCRP Report No. 11/2018, Geneva

Wold, S., Ruhe, A., Wold, H., and Dunn, W.J. (1984). The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM J. on Scientific and Statistical Computing. 5 (3), 735–743, doi:10.1137/0905052

World Bank (2014). “Financial Protection Against Natural Disasters: From Products to Comprehensive Strategies.” Washington, D.C.: The World Bank.

WRI – World Resources Institute (2020). “New data shows millions pf people, trillions in property at risk from flooding – but infrastructure investments now can significantly lower flood risk. Press Release, 23 April 2020, <https://www.wri.org/news/release-new-data-shows-millions-people-trillions-property-risk-flooding-infrastructure>

Yami, M., Meyer, F., and Hassan, R. (2020). The impact of production shocks on maize markets in Ethiopia: implications for regional trade and food security. Agricultural and Food Economics, 8, doi:10.1186/s40100-020-0153-5

Yang, W., Seager, R., Cane, M. A., and Lyon, B. (2015). The annual cycle of East African precipitation. J. Clim., 28(6), 2385–2404, doi:10.1175/JCLI-D-14-00484.1

Yang, M., Wang, G., Ahmed, K. F., Adugna, B., Eggen, M., Atsbeha, E., You, L., Koo, J., and Anagnostou, E. (2020). The role of climate in the trend and variability of Ethiopia's cereal crop yields. *Science of the Total Environment*, 723, doi:10.1016/j.scitotenv.2020.137893

Yang, M., Wang, G., Lazin, R., Shen, X., and Anagnostou, E. (2021). Impact of planting time soil moisture on cereal crop yield in the Upper Blue Nile Basin: a novel insight towards agricultural water management. *Agricultural Water Management*, 243, doi:10.1016/j.agwat.2020.106430

Yesuf, M., and Bluffstone, R. (2007). Risk aversion in low income countries: Experimental evidence from Ethiopia. IFPRI Discussion Paper No. 00715, doi:10.22004/ag.econ.42398

Yu, X., and Tang, Y. (2017). A critical review on the economics of disasters. *J. Risk Analysis and Crisis Response*, 7(1), 27-36, doi:10.2991/jrarc.2017.7.1.4

Zhang, Y., Moges, S., and Block, P. (2016). Optimal cluster analysis for objective regionalization of seasonal precipitation in regions of high-spatial-temporal variability: Application to Western Ethiopia. *Journal of Climate*, 29, 3697-3717

Zhang, Y., You, L., Lee, D., and Block, P. (2020). Integrating climate prediction and regionalization into an agro-economic model to guide agricultural planning. *Climatic Change*, 158, doi:10.1007/s10584-019-02559-7

ProQuest Number: 28965510

INFORMATION TO ALL USERS

The quality and completeness of this reproduction is dependent on the quality
and completeness of the copy made available to ProQuest.



Distributed by ProQuest LLC (2022).

Copyright of the Dissertation is held by the Author unless otherwise noted.

This work may be used in accordance with the terms of the Creative Commons license
or other rights statement, as indicated in the copyright statement or in the metadata
associated with this work. Unless otherwise specified in the copyright statement
or the metadata, all rights are reserved by the copyright holder.

This work is protected against unauthorized copying under Title 17,
United States Code and other applicable copyright laws.

Microform Edition where available © ProQuest LLC. No reproduction or digitization
of the Microform Edition is authorized without permission of ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346 USA