Title
Impacts of climate variability, decision-making and digital information on agricultural outcomes in sub-Saharan Africa

Author
Krell, Natasha

Publication Date
2021

Peer reviewed|Thesis/dissertation

University of California
Santa Barbara

Impacts of climate variability, decision-making and digital information on agricultural outcomes in sub-Saharan Africa

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

Doctor of Philosophy
in
Geography

by

Natasha T. Krell

Committee in charge:

Professor Kelly Caylor, Chair
Professor Chris Funk
Professor Stuart Sweeney

June 2021

The Dissertation of Natasha T. Krell is approved.

Professor Chris Funk

Professor Stuart Sweeney

Professor Kelly Caylor, Committee Chair

June 2021

Impacts of climate variability, decision-making and digital information on agricultural outcomes in sub-Saharan Africa

by

Natasha T. Krell

To James Hamilton Pickens and Ian Akeem Lemaiyan

Acknowledgements

I would first like to thank my advisor Kelly Caylor for his kind mentorship and guidance; my committee members Chris Funk and Stuart Sweeney for their support and expertise; and my co-authors for their contributions: Sara Lopus, Stacey Giroux, Corrie Hannah, Zack Guido, Tom Evans, Bryn Morgan and Drew Gower. I am grateful for my friends and labmates who have been hugely supportive and offered feedback: Keita DeCarlo, Marc Mayes, Rachel Green, Sophia Arabadjis, Kate Voss, Chelsie McWhorter, Emily Williams, Katelyn Zigner, Sheila Saia and Kevin Roche.

I am grateful for my partner Josh Novins who supported me throughout this journey and for my parents and brother who give me the freedom to chart my own course. I am also grateful for my Santa Barbara family, Kelly and Federico Peinado, who will always hold a special place in my heart when I look back on my graduate school years.

Finally, I would like to acknowledge the farmers who participated in this research and shared their stories and insights with me and our field staff at the Mpala Research Center and the Zambia Agriculture Research Institute: John Gitonga, Boniface Kimathi, Eligah Mutanda, and Allan Chillenga for facilitating this research.

Curriculum Vitæ
Natasha T. Krell

Education

2021	Ph.D. in Geography, University of California, Santa Barbara (UCSB) Emphasis in Climate Sciences and Climate Change
2016	B.A. in Human Ecology, College of the Atlantic (COA)

Fellowships, Scholarships, Grants and Awards

2021	Excellence in Research Award, *UCSB Dept. of Geography* Chair's Excellence in Service Award, *UCSB Dept. of Geography*
2020	Summer Graduate Fellowship, *UCSB Earth Research Institute* Coastal Fund Grant, Bird-window Collisions Working Group *UCSB*
2019	Environmental Solutions Fellowship, *Schmidt Family Foundation* Science, Mathematics, and Research for Transformation Scholarship, *Department of Defense*
2018	Finalist for Fulbright Research Fellowship to Kenya, *Institute of International Education* Graduate Scholars Program, *UCSB Graduate Division* NASA Remote Sensing Hydrology Workshop Student Travel Grant, *Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI)*
2016	Maine Space Grant Consortium Award, *NASA*
2015	Shelby Cullom Davis International Advanced Studies Award, *COA* Kathryn W. Davis Global and Civic Engagement Fund Award, *COA* Maine Space Grant Consortium Award, *NASA* Presidential Scholarship, *COA* Sierra Club Coalition SPROG Scholarship, *COA*
2014	International Student Travel Grant, *American Geophysical Union* Rothschild Student-Faculty Collaboration Grant, *COA* Presidential Scholarship, *COA* Sierra Club Coalition SPROG Scholarship, *COA*

2013	Len Assante Scholarship, *National Groundwater Association* Environmental Leadership Grant, *Henry David Thoreau Foundation*

Publications

In review	**Krell, N.,** Davenport, F., Harrison, L., Turner, W., Peterson, S., Shukla, S., Husak, G., Evans, T,. and Caylor, K. Using real-time mobile phone data to characterize the relationships between small-scale farmers' planting dates and earth observation datasets. Climate Risk Management.
	Krell, N., Morgan, B., Gower, D., and Caylor, K. Consequences of dryland maize planting decisions under increased seasonal rainfall variability. Water Resources Research.
	Guido, Z., Lopus, S., Waldman, K., Hannah, C., Zimmer, A., **Krell, N.,** Knudson, C., Estes, L., Caylor, K., and Evans, T. Perceived Links Between Climate Change and Weather Forecast Accuracy: New Barriers to Tools for Agricultural Decision-making. Climatic Change.
	Joshi, N., Gerlak, A., Hannah, C., Lopus, S., **Krell, N.,** Evans, T. Water insecurity, housing tenure and the role informal water providers in Nairobi's slum settlements. World Development.
	Davenport, F., Shukla, S., Turner, W., Harrison, L., **Krell, N.,** Husak, G., Lee, D., Peterson, S. Sending out an S.O.S: Using Start of Rainy Season Indicators for Market Price Forecasting to Support Famine Early Warning. Environmental Research Letters.
2021	Shukla, S., Husak, G., Turner, W., Davenport, D., Funk, C., Harrison, L., **Krell, N.** A slow rainy season onset is a reliable harbinger of drought in most food insecure regions in Sub-Saharan Africa. Plos one. https://doi.org/10.1371/journal.pone.0242883.
	Hannah, C., Giroux, S., **Krell, N.,**, Lopus, S.,McCann, L., Zimmer, A., Caylor, K., Evans, T. Has the vision of a gender quota rule been realized for community-based water management committees in Kenya? World Development. https://doi.org/10.1016/j.worlddev.2020.105154.

2020 Guido, Z., Zimmer, A., Lopus, S., Hannah, C., Gower, D., Waldman, K., **Krell, N.,** Sheffield, J., Caylor, K., Evans, T. Farmer Forecasts: Impacts of Seasonal Rainfall Expectations on Agricultural Decision-Making in Sub-Saharan Africa. Climate Risk Management. https://doi.org/10.1016/j.crm.2020.100247.

Krell, N., Giroux, S.A., Guido, Z., Hannah, C., Lopus, S.E., Caylor, K.K. and Evans, T.P., 2020. Smallholder farmers' use of mobile phone services in central Kenya. Climate and Development, pp.1-13. https://doi.org/10.1080/17565529.2020.1748847

2016 Boyd, R. S., **Krell, N.,** and Rajakaruna, N. Extreme Environments. In: Oxford Bibliographies in Ecology Ed. David Gibson. New York: Oxford University Press.

Peer-Reviewed Oral Presentations

2020 **Krell, N.,** Morgan, B., Gower, D., and Caylor, K. K. "Consequences of variable rainfall on farming outcomes for dryland maize farmers." American Geophysical Union Fall Meeting. Virtual.

2017 **Krell, N.,** Evans T.P., Estes, L.D., and Caylor, K.K. "Real-time monitoring of smallholder farmer responses to intra-seasonal climate variability in central Kenya." American Geophysical Union Fall Meeting. New Orleans, LA.

Invited Talks

2020 **Krell, N.** "An Ecohydrological Model for Simulating Rainfall Variability Impacts on Maize Production." UCSB Graduate Simulation Seminar Series (GS3). Virtual.

Krell, N. "Farming Fast and Slow: Episode 9: Infield Weather Data and Smallholder Farmers." Arable Labs. Online Webinar.

2019 **Krell, N.** "Use of mobile phones for agriculture in Kenya." Lightning Talk. Schmidt Environmental Solutions Fellows Open House at UCSB. Santa Barbara, CA.

Krell, N. "To what extent does climate variability explain farmers' planting decisions in central Kenya?" Graduate associate lunch talk for UCSB Broom Center for Demography. Santa Barbara, CA.

Krell, N. "Muddling through with muddy boots: Conducting Fulbright research in Kenya." Invited talk for Climate Change Seminar Series. College of the Atlantic, Bar Harbor, ME.

Evans, T. (presenting author), Lopus, S., Guido, Z., Zimmer, A., Hannah, C., Dell'Angelo, Caylor, K., Tuholske, C., **Krell, N.**, and Estes, L. "Perceptions of population growth vs. climate change as threats to irrigated agriculture in Kenya." Broom Center for Demography Seminar Series. University of California, Santa Barbara.

2018 **Krell, N.** "Gender differences in access/use of mobile phones, agricultural management, and engagement in farmer cooperatives in central Kenya." Graduate associate lunch talk for UCSB Broom Center for Demography. Santa Barbara, CA.

Peer-Reviewed Poster Presentations

2020 **Krell, N.** "Consequences of increasing rainfall variability on farming outcomes in dryland maize systems." Graduate Climate Conference. Virtual.

2019 **Krell, N.,** Davenport, F, Peterson, S., Shukla, S., Husak, G.J, Turner, W., Funk, C.C., Caylor, K.K., "To What Extent Does Climate Variability Explain Farmers? Planting Decisions in Central Kenya?" American Geophysical Union Fall Meeting. San Francisco, CA.

Krell, N. "When Mentee Becomes Mentor: Graduate Perspectives on Mentorship." American Geophysical Union Fall Meeting. San Francisco, CA.

2015 **Krell, N.,** DeCarlo, K.F., and Caylor, K.K. "Analysis of Biophysical Mechanisms of Gilgai Microrelief Formation Using Ultra-High Resolution Aerial Imagery." American Geophysical Union Fall Meeting. 2015. San Francisco, CA.

Krell, N., Dawson, H.R, and Rajakaruna, N. "Edaphic-climatic influences on the ecology and evolution of plants found on serpentine and granite outcrops of Deer Isle, Maine." Northeast Natural History Conference. Springfield, MA.

2014 **Krell, N.,** Papuga, S.A., Kipnis, E., Nelson, K. "Dynamic Pulse-Driven Flowering Phenology in a Semiarid Shrubland." American Geophysical Union Fall Meeting. San Francisco, CA.

Krell, N., Papuga, S.A., Kipnis, E., Nelson, K. "Dynamic Pulse-Driven Flowering Phenology in a Semiarid Shrubland." Research Insights in Semiarid Ecosystems (RISE) Symposium. Tucson, AZ.

Krell, N., Papuga, S.A., Kipnis, E., Nelson, K. "Dynamic Pulse-Driven Flowering Phenology in a Semiarid Shrubland." Phenology Research and Observations of Southwest Ecosystems (PROSE) Symposium. Tucson, AZ.

Research Positions

2019-2021 **Graduate Student Researcher,** Climate Hazards Center
Mentors: Dr. Chris Funk, Dr. Frank Davenport

2016-2019 **Graduate Student Researcher,** Earth Research Institute
Mentors: Dr. Kelly Caylor & Dr. Tom Evans

2015 **Research Internship,** Mpala Research Centre, Laikipia, Kenya
Mentor: Dr. Kelly Caylor

2014-2016 **Independent Research,** College of the Atlantic, USA
Mentor: Dr. Nishanta Rajakaruna

2015 **Research Assistant,** Acadia National Park, USA
Mentor: Caitlin McDonough MacKenzie

2014 **NSF-REU Internship,** University of Arizona, USAl
Mentor: Dr. Shirley (Kurc) Papuga

2012-2013 **Research Assistant,** Smith College, USA
Mentor: Dr. Andrew Guswa

Teaching

Spring 2020 **Certificate,** Pillars of Teaching Assistantship Certificate
Completed UCSB Instructional Development's Pillars of Teaching

Assistantships workshops for training in effective and research-based teaching practices.

Fall 2017 **Teaching Assistant,** GEOG 3A, U.C. Santa Barbara, USA
Led three discussion sections for introductory undergraduate Geography course: Oceans and Atmosphere. Professor: Ms. Tessa Montini.

Summer 2017 **Teaching Assistant,** EEMB 120, U.C. Santa Barbara, USA
Led three discussion sections for undergraduate course in the Department of Ecology, Evolution, and Marine Biology: Intro to Ecology. Professor: Dr. Hillary Young.

Spring 2017 **Teaching Assistant,** GEOG 167, U.C. Santa Barbara, USA
Led discussions and activities for upper-division undergraduate Biogeography class at UCSB (GEOG / ENV S 167). Professor: Dr. Kelly Caylor.

Community Outreach and Service

2020-2021 **Graduate Mentor,** Bird-window Collisions Working Group
Managed four undergraduate interns and major Coastal Fund grant as part of student-led initiative to reduce bird-window collisions on campus.

2019-2021 **Representative,** Chair's Graduate Advisory Committee
Selected to be a graduate student representative on the Advisory Committee to the Chair of the Department of Geography at UCSB.

2019-2020 **Graduate Mentor,** Graduate Scholars Program
Mentored three first-year graduate students on their academic progress, career development, and general adjustment to graduate school.

2017-2020 **Rising Stars Mentor,** Wedu Global Foundation
Mentored Nepali women college students via Skype once a month. Provided guidance, support, and lead mentees in personal and professional development activities.

2019 **Graduate Mentor,** PIPELINES Program
Mentored three undergraduate students on an independent research project using unmanned aerial systems for environmental monitor-

ing and image processing as part of NAVFAC EXWC summer mentorship program in partnership with UCSB's Center for Science and Engineering Partnerships (CSEP).

2016-2019 **Technical Committee on Ecohydrology,** AGU
Served as student representative on AGU Ecohydrology Technical Committee.

2015-2016 **Hydrology Section Student Subcommittee,** AGU
Elected for two-year position on the American Geophysical Union's Hydrology Section Student Subcommittee. Organizer of 2015 Student Conference and co-convener of 2015 and 2016 Social Dimensions of Geoscience pop-up talks.

2014-2015 **Admissions Committee,** College of the Atlantic
Served as undergraduate representative on admissions committee at College of the Atlantic. Reviewed applications to admit transfer and first-year students to College of the Atlantic's class of 2019.

Languages and Software

Spoken Spanish (fluent), Kiswahili (advanced), Mandarin (intermediate)

Computational Python, R, ArcGIS, QGIS, MATLAB, Git, LaTeX

Abstract

Impacts of climate variability, decision-making and digital information on agricultural outcomes in sub-Saharan Africa

by

Natasha T. Krell

Climate variability poses significant threats to agricultural production, particularly for small-scale and rainfed farmers who are often the first to feel the impacts of extreme weather and are the most vulnerable to climate change. Timely information disseminated via mobile phones is one pathway to manage risk and reduce vulnerabilities to a changing climate. However, poor service coverage, technological inequities and lack of awareness about mobile phone-based services pose significant barriers to access and use of information. In my dissertation, I asked questions about the impacts of climate variability on small-scale farming systems and investigated the roles of information communications technologies (ICTs) in the diffusion of digital-based information.

This dissertation focuses on small-scale agriculturalists in central Kenya and on telecommunications access in Zambia. In Chapter 2, I conducted a household survey to report on the use of mobile phones for agriculture in central Kenya. I found that while most respondents owned a mobile phone, the overall percentage of adoption of mobile phone services for agriculture is low. Leveraging farmer groups is a potential path forward in the spread of digital information. In Chapter 3, I developed an ecohydrological model and simulated rainfall to model crop water stress for maize. I highlighted nonlinearities between season rainfall and yield that led to divergent outcomes for farmers. Given the impacts of climate variability which are already being realized, there is an opportunity to disseminate real-time alerts via mobile phones. In Chapter 4, I returned

to the question of access to telecommunications and applied a Gibbs point process model to the location of cell towers in Zambia. I found persistent disparities in telecommunications access between rural and urban residents and remote regions disconnected from road networks. These three studies provided key insights into the changing landscape of small-scale farming, and offered practical guidance for improving dissemination and uptake of targeted and mobile-based agro-meteorological information, services and alerts.

Contents

Chapter 1

Introduction

While small-scale farmers have historically had to manage fluctuations in weather and year-to-year variability, they now must deal with unprecedented and erratic climate patterns (Morton, 2007) which are likely to be exacerbated by climate change (IPCC, 2014). Rising temperatures and shifting patterns of rainfall are projected to negatively impact agricultural output (Niles and Brown, 2017; Funk and Brown, 2009; Lobell et al., 2011; Ray et al., 2019), particularly in sub-Saharan Africa (SSA) where the majority of land is arid and semiarid, and smallholder systems dominate agricultural production (Muller et al., 2011). Small-scale farms produce the majority of the world's food and yet they only make up 12% of global farmland (Ricciardi et al., 2021). These farmers face growing pressure of climate-induced losses in production (Ricciardi et al., 2021).

Sub-Saharan Africa (SSA) is the most vulnerable region to climate variability and change (Slingo et al., 2005) due to several factors linked to its physical geography (e.g. already high temperatures) and reduced adaptive capacity (Downing et al., 1997; McCarthy et al., 2011). SSA agriculture is highly sensitive to climate shocks (e.g. spikes in temperature and precipitation, and extreme weather events), is characterized by lack of irrigation infrastructure, limited access to inputs (e.g. high yield seed varieties, fertilizers, etc.) (Becker-Reshef et al., 2020) and has a sizeable rural population dependent on agricultural livelihoods and domestic economies (Dinar et al., 2008). The combination of these factors make climate change adaptation challenging in SSA. Low yield variance is important for stable farmer incomes (Reidsma et al., 2010), reliable food supply (Lobell et al., 2011; Slingo et al., 2005), and preventing spikes in food prices (Ray et al., 2012). To some extent, African farmers are already adapting to climate variability (Kurukulasuriya et al., 2006). Increasing temperatures, altered rainfall patterns, and extreme weather such as droughts and floods, threaten SSA agricultural production and may force marginal land out of production (Kurukulasuriya et al., 2006).

One path forward for improving adaptation of agriculture is through the dissemination of crop advisories, seasonal rainfall forecasts, and early warning alerts. In a report by the Global Commission on Adaptation, it was found that strengthening Early Warning Systems is an area with a high benefit-cost ratio (10:1) compared to improving resilience of new infrastructure, making water resources management more resilient or improving dryland agriculture crop production (Bapna et al., 2019). Satellites and remote sensing technologies provide us with Earth observation (EO) datasets that provide critical and life-saving information and data (Nakalembe, 2020). This information can aid early warning about natural hazards such as droughts (Nakalembe, 2018), flooding (Kilavi et al., 2018) and pest outbreaks such as locusts (Cressman, 2013; Nakalembe, 2020). Typically, this information has focused on rescue operations after natural disasters occur. However, additional resources need to be allocated to improve monitoring and early warning systems in order to be proactive about such events (Nakalembe, 2020).

Mobile phones can be a powerful tool in the dissemination of potentially life-saving information in the face of climate hazards. Information accessed by African farmers via mobile technologies continues to rise (Baumüller, 2018) with the exponential increase in mobile phone ownership since the early 2000s (World Bank, 2018a). Cellular networks enable individuals to communicate and collaborate via information communication technologies (ICTs) and provide access to financial services such as mobile banking. Increasingly, real-time information on extreme weather, pest and disease outbreaks, and other agro-climatic services are disseminated through ICT devices, including smartphones (Nakasone et al., 2014; Baumüller, 2018). However due to systematic and prevailing inequalities, women and other marginalized groups are often left without access to ICTs (UN Women, 2015).

Overall, this dissertation seeks to advance the science on climate impacts on rainfed maize cropping systems and the landscape that would be make early warning systems accessible and usable at the household and community levels. The purpose of an early warning system is to prompt an early response before a crisis manifests which then requires substantial resources to remediate (Becker-Reshef et al., 2020). Given that climate change is not going away anytime soon, funds and research endeavors need to focus on proactive warning rather than only dealing with the consequences after the fact.

I conducted research in two African countries: Zambia and Kenya. Kenya is a leader in innovation and use of ICTs and Kenyan smallholder agriculture is highly susceptible to spatio-temporal fluctuations in rainfall. With the majority of its population residing in rural areas, Zambia poses a challenge to widespread coverage of ICTs in part due to its sparse population spread diffusely across the country. This dissertation focuses on the following overarching questions: (1) What are the impacts of rainfall variability on small-scale production in rainfed agricultural systems and what management strategies can be employed to minimize crop failure? (2) How does access to and use of information communication technologies (ICTs) vary at community- and national-scales; specifically, who uses climate information services for agriculture? (3) What is the spatial distribution of cell towers in Zambia?

1.1 Permissions and Attributions

1. Chapter 2 is an Accepted Manuscript of an article published by Taylor & Francis in *Climate and Development* on 9 April 2020, available online[1]. The article was in collaboration with Dr. Stacey Giroux (Indiana University), Dr. Zack Guido (University of Arizona), Dr. Corrie Hannah (University of Arizona), Dr. Sara Lopus (California Polytechnic State University), Dr. Kelly Caylor (University of California, Santa Barbara), and Dr. Tom Evans (University of Arizona). It is reproduced here with the permission of Taylor & Francis.

2. Chapter 3 is a manuscript currently under review in the journal *Water Resources Research*. The article was in collaboration with Bryn Morgan (University of California, Santa Barbara), Dr. Drew Gower (University of Maryland), and Dr. Kelly Caylor (University of California, Santa Barbara).

[1] https://www.tandfonline.com/doi/full/10.1080/17565529.2020.1748847

Chapter 2

Smallholder farmers use of mobile phone services in central Kenya

Abstract Information and services delivered through mobile phones, "m-services", have transformative potential to provide rural African farmers with important agro-meteorological information. However, a greater understanding is needed regarding the types of m-services available to farmers, how farmers access that information, and possible factors affecting m-service use. With a study of smallholder farming communities in central Kenya, we examine factors affecting the likelihood of Kenyan farmers' adoption of m-services specifically related to agriculture and livestock information, buying and selling products, and alerts about agricultural or livestock activities. According to a survey of 577 farming households, 98% of respondents own a mobile phone. Approximately 25% use it to access information about agriculture and livestock, 23% access information about buying and selling products, and 18% receive alerts. Personal smartphone ownership increases the likelihood of m-services use ($p < 0.001$), as does membership in farmer organizations ($p < 0.001$). We find that age and income are not significantly related to m-service use, and we discuss this result in terms of intersections between smartphone

ownership, gender, and education. Consistent with prior qualitative research, our quantitative results further support the need for m-services providers to design m-services for basic or feature phone users. Understanding how gender and broad socioeconomic differences shape farmers' access to and use of mobile phones for agriculture provides baselines about diverse sets of users, which can help inform efficient and targeted dissemination of agro-meteorological information services[1].

2.1 Introduction

Lack of access to information and knowledge transfer can hamper agricultural production in rural farming communities in sub-Saharan Africa (SSA). Agricultural, market, and weather information is critical to agricultural productivity, especially for reducing uncertainty and risk associated with extreme weather events and disease (Baumüller, 2013). The dissemination of agro-meteorological information can improve livelihoods by reducing uncertainty and enable improved inputs and technology adoption (Hansen et al., 2007). Access to information through mobile phones and mobile internet can also help agriculturalists manage risk and reduce vulnerabilities to a changing climate (Baumüller, 2013).

Information communication technologies (ICTs) such as mobile phones are touted as digital platforms with transformative potential to reach many farmers at once across rural settings (World Bank, 2018a; Santosham and Lindsey, 2015). As the cost of mobile phones have fallen and connectivity has spread, phone ownership and internet access have become possible for populations in the continent's lowest-income areas (Wyche and Olson, 2018). With this uptake of mobile phones, users can subscribe to receive mobile

[1]This is an Accepted Manuscript of an article published by Taylor & Francis in *Climate and Development* on 9 April 2020, available online: online `https://www.tandfonline.com/doi/full/10.1080/17565529.2020.1748847`.

phone-enabled services or "m-services" to access agro-meteorological (Baumüller, 2013) and market information (Wyche and Steinfield, 2016).

M-services deliver electronic media content through mobile technologies and is an umbrella term that includes m-agri, m-commerce, m-banking or m-payments. M-services come in varied forms, including Short Message Service (SMS), Unstructured Supplementary Service Data (USSD), mobile applications (apps) and helplines. The difference between SMS and USSD protocol is that SMS is a text messaging service whereas USSD protocol are in the form of "Quick Codes". Depending on the electronic media m-services contain, they can be accessed by phones with and without internet access. M-services can be used to connect buyers to sellers, disseminate general information about farming and livestock (such as market information on prices), and send alerts on pest and disease threats (Baumüller, 2018). Some m-services are free to use or may require a cost to use advanced features, while others are entirely proprietary. For example, Ujuzi Kilimo in Kenya offers actionable recommendations to farmers through subscription-based SMS and USSD services[2].

Whether m-services can improve agricultural livelihoods is a question facing scholars and development programs focused on addressing rural livelihood vulnerability. Qiang et al. (2012) showed that increased access to climate, crop disease, and market information via m-services improved farmers' production and profitability in Kenya. However, wealthier, educated, and typically urban populations have greater access and therefore benefit from m-services in comparison to rural, poorer populations, especially rural women (Souter et al., 2005; Porter et al., 2012; Wyche and Olson, 2018; Wyche et al., 2019). As a result of this limited access to information for some populations, scholars have questioned whether mobile-based market information can improve circulation of market prices and reduce information asymmetries between farmers and buyers. Srini-

[2]https://www.ujuzikilimo.com/sms.html

vasan and Burrell (2013) and Wyche and Steinfield (2016) have detailed the underlying barriers to using mobile phones for accessing Market Information Systems (MIS). These barriers include, but are not limited to, cost of airtime, challenges with charging faulty and low-quality batteries, language, and literacy (Srinivasan and Burrell, 2013; Wyche and Steinfield, 2016).

However, in the literature on ICT use in SSA, far less attention has been placed on understanding the different types of information communicated via m-services, how farmers access that information, and possible factors affecting the likelihood of m-service use. Our study addresses these knowledge gaps by drawing on a sample of more than 500 smallholder farmers in rural central Kenya. We first identify who uses m-services, the types of m-services that survey respondents use, and how the m-services are used. We then assess the factors that affect the likelihood of m-service use by modeling the associations between individual-level characteristics and three classes of agricultural information available via m-services: farming information, buying and selling farming products, and alerts on farming activities. The latter helps us understand the underlying factors impeding m-services adoption. We then discuss the interactions between education, income, and gender with smartphone ownership; the important role that farmer organizations play in m-services adoption; and how developers of m-services can use this information to target unreached individuals.

2.2 Literature review

2.2.1 ICTs, adaptive capacity, and vulnerability

ICTs are often cast as technologies that can increase access to information and resources and connect individuals. Information disseminated via m-services are therefore seen as important tools for helping farmers adapt and to address vulnerability (Eakin et al., 2015), where vulnerability, per the IPCC (2007) is "the degree to which an environmental or social system is susceptible to, and unable to cope with, adverse effects of climate change, including climate variability and extremes" (p. 883). Vulnerability is often conceptualized as including three interlinking elements: exposure, sensitivity, and adaptive capacity (Adger, 2006). While exposure relates to the degree and type of the perturbation, adaptive capacity relates to the capacity of individuals or groups to manage and influence their resources and risks in the face of a perturbation (Waters and Adger, 2017). Both the degree of exposure and adaptive capacity shape a system's sensitivity to that perturbation. Many determinants of household-level adaptive capacity have been identified, which can relate to access to assets and resources (Moser, 1998). At the local level, one such determinant has been access to information resources and the ability of decision-makers to marshall the information (Smit and Wandel, 2006; Fawcett et al., 2017). Within this context, ICTs provide an efficient means to reach a growing user base and build adaptive capacity through enabling access to critical information and facilitating a process of learning (Eakin et al., 2015).

There is, however, contradictory evidence about how useful ICTs are or can be for addressing vulnerability, with some studies finding positive effects and others no impact. In a review of climate change adaptation and ICTs in the Caribbean and Latin America, Eakin et al. (2015) suggest that ICTs support adaptation through increasing social capital, improving access to critical information for decision-making, and coordinating actors. Marenya and Barrett (2007) found that lack of access to information, land, and credit, constrained natural resource management efforts and thus rendered smallholder farmers in western Kenya more vulnerable to climate variability. Similarly, researchers contended that in Laikipia county, Kenya, limited access to agro-meteorological information hampers adaptive capacity (Wiesmann, 1998; Ogalleh et al., 2012). Bryan et al.

(2009) came to similar conclusions that lack of access to information is a main barrier to adaptation among Ethiopian farmers. Specific to the impact of DrumNet, a phone-based MIS, Ogutu et al. (2014) found a positive influence on labor productivity, seeds, fertilizers and land in three Kenyan provinces.

A growing body of literature criticizes ICTs for development programs and provides evidence for why they do not work. In the domain of distributing market price information through MIS, Camacho and Conover (2011) and Fafchamps and Minten (2012) showed poor adoption of MIS and their lack of impact on agricultural outcomes. Following this work, Burrell and Oreglia (2015) sought to better understand Ugandan and Chinese agriculturalists' decision-making processes and need for market price information. They suggested that using mobile phones to collect and distribute market information through MIS is of limited relevance since information about price is only one of several factors that aids decision-making (Burrell and Oreglia, 2015). Similarly, Srinivasan and Burrell (2013) suggest that mobile phones should not be given an over-privileged role in seeking market price information. Finally, although there are many apps for agriculture in Kenya and SSA, it is still unclear how many people actually use them after initially subscribing and/or downloading them.

2.2.2 Adoption and utilization of m-services in Kenya

Adoption of mobile phones and m-services is attributed to several factors related to the local context. In Kenya, agriculture and livestock are key economic sectors and contribute to more than one third of total gross domestic product (GDP) (Kenya Bureau of Statistics Economic Survey, 2017). The Communications Authority of Kenya (2018) reports that as of March 2018, 95.1 percent of the adult population had a mobile phone subscription and 42.9 percent had access to broadband internet. The growing affordability of internet access coincides with an impressive number of m-services available for Kenyan farmers to utilize. For instance, Twiga Foods connects horticultural farmers to buyers using a mobile-based buyer-to-buyer platform[3], and Mkulima Online, M-Farm, provides access to crop prices and connects buyers and sellers (Baumüller, 2013; Baumüller, 2015). The high adoption rates of mobile phones in Kenya compared to other sub-Saharan African countries, see Figure 2.1, may be one factor for the multitude of m-services available.

Given the ubiquity of mobile phone ownership in Kenya, ICTs may continue to have increased importance in the sphere of agricultural extension. The aim of agriculture extension is to provide services and advice to rural farmers and their families so they may maximize the resources made available to them (Katz and Barandun, 2002). Private extension services exist in Kenya in part due to the inefficiency of public extension, which was identified as a factor that impedes agricultural development and perhaps an explanation for low yields (Muyanga and Jayne, 2008). M-services can serve as a gap-filling mechanism for the agriculture extension system. While governments may provide their own m-services through websites and/or SMS or USSD services, partnerships with other sectors such as private companies, NGOs, and research institutions can help bolster farmer productivity and improve the limited capacity of government programs (Caine et al., 2015; Donovan, 2017).

2.2.3 Factors that influence adoption and use of M-services

Factors that influence technology adoption and use are typically related to education level, age, and gender (Meso et al., 2005); however cost may also pose a barrier to usage. Ogutu et al. (2014) found a significant difference in average age between DrumNet participants and nonparticipants; however no difference in gender. Wyche and Steinfield

[3] https://twiga.ke

(2016) investigated factors that impede adoption of the m-service M-Farm in western Kenya. The majority of interviewees owned feature phones, and while participants could theoretically access M-Farm with those phones, they did not. Wyche and Steinfield (2016) note a variety of barriers including limited amounts of phone credit which stifles SMS use, phone charging limits, phones with considerable wear and tear that prohibited use, as well as the perception that mobile phones are for voice communication rather than SMS-interaction.

After initial adoption of m-services, appropriate use of information requires having credible information and trusting in that information. Mittal et al. (2010) found that farmers use mobile-enabled agricultural information when the information is timely, of good quality, and when they trust the information. For SMS information specifically, trust needs to be established: the recipient is unlikely to take SMS information coming from an unknown sender in the slurry of other spam SMS (Crandall, 2012; Cheney, 2018).

Research also suggests that individuals' perceptions are important determinants of individual technology use. Thiga and Ndungu (2015), for example, identify lack of awareness as the primary reason why agriculture extension officers do not utilize ICTs in Kenya. Mobile applications in particular were the least utilized form of ICT among respondents (Thiga and Ndungu, 2015). Since agriculture extension officers provide outreach to farmer organizations such as agricultural and livestock cooperatives they are positioned to be agents of change and advancement within a community. However, in cases where the extension agents are not exhibiting high adoption or promotion of these useful services, farmers may benefit from information transfer through other routes such as mobile devices and their non-extension social networks.

2.3 Research methodology

2.3.1 Study area

Laikipia, Meru and Nyeri counties meet at the northern and western slopes of Mount Kenya in the semiarid highlands of Kenya. We focus on 35 smallholder farming communities within the following sub-counties: Laikipia North and Laikipia East (Laikipia), Buuri, North and Central (Meru), and Kieni East (Nyeri). Many of the households are part of farmer organizations, which can be classified as farmer cooperatives, farmer groups, or local water resource governance groups called Community Water Projects (CWPs). In this study, we characterize farmer cooperatives and CWPs as formal organizations in comparison to farmer groups, which are informal organizations.

Farmer cooperatives are official, government-registered forms of collective action that are typically composed of smallholder farmers working together across a large area on agricultural production, sale of products, and opportunities to enter higher-value markets (Narrod et al., 2009; Markelova and Mwangi, 2010). They follow a formal structure of organization with written rules. One of the main benefits of farmer cooperative memberships is the reduction of transaction costs among smallholder agriculture producers who are often located in remote areas and have limited economic capacity to enter production systems (Markelova and Mwangi, 2010). In contrast to farmer cooperatives, farmer groups are informal and may not be officially registered with the government. Farmers either self-organize or are organized into groups by agricultural extension agents to facilitate trainings and knowledge exchange. Members in these informal organizations typically share information about best practices in agriculture. They may also collectively pool savings to help afford agricultural inputs or assist group members in times of need. Community Water Projects (CWPs) are a third type of farmer organization. CWPs use formal institutions where members officially meet and interact with each other on a weekly to monthly basis to attend meetings, maintain irrigation infrastructures, and access irrigation water resources. To become a CWP member, farmers typically pay a joining fee to connect to irrigation water via piped networks from rivers off Mount Kenya, as well as monthly maintenance fees.

2.3.2 Data collection

A team of eight Kenyan enumerators conducted the household survey using Qualtrics (2019) software between June and July 2018 (Qualtrics, Provo, UT). The multilingual team of six women and two men conducted the survey in Kiswahili, Kikuyu, or Kimeru, depending on the respondent's background. We interviewed 605 respondents; however 577 responses were used for the study. We removed respondents who either did not indicate farming as their primary occupation, were flagged by enumerators as giving incomplete answers, or refused to answer or did not know their level of education attained. Respondents were not compensated for their participation in the study. See Supplementary Info (S.I.) for more information on survey methodology[4].

The households selected for the study are a representative sample of households that receive water from CWPs. Additionally, our sample is not representative of all Kenyan smallholders. These households were selected as part of a five-year multi-institutional research project conducted in the region (McCord et al., 2017; Lopus et al., 2017). The five-year study assesses agronomic decision-making of irrigated (CWPs) and non-irrigated households in Laikipia, Meru and Nyeri counties. At the start of the project, we used a randomized sampling approach of farmers within CWPs. As the project expanded, we gathered longitudinal data as well as data from non-CWP members because we could not get in contact with everyone who had previously been sampled. Thus, we augmented our dataset with the help of local guides and selected neighboring households within the

[4]https://doi.org/10.6084/m9.figshare.12107166.v1

communities. Respondents were either the head of household or spouse of the household head.

The survey took 90 minutes on average and covered a range of topics including agricultural management, perceptions of rainfall and climate change, use of weather and climate services, use and barriers to use of ICTs, migration, and household socio-demographics. The ICT module was administered about halfway through the survey and took approximately 20 minutes. The ICT module followed the design of the U.S. Agency for International Development's (USAID) toolkit on gender and ICT (Highet et al., 2017) and comprised of three parts: demographics, access and use of ICTs, and mobile farming use.

The USAID practical toolkit offers quantitative and qualitative methods for data collection on access, usage, barriers and perceptions of ICTs such as mobile phones, radios and other internet-enabled devices. We selected and modified questions from the toolkit and we pre-tested the questions in three pilot interviews. Additionally we used a four-day enumerator training session to refine the questionnaire; however no major corrections were made. Our analysis focuses on use and non-use of three m-services, which was determined by the responses to the question: "During the last growing season, did you use your mobile phone to access any of the following services for your agricultural/livestock management? (e.g. includes Facebook, Mkulima Bora, WhatsApp, M-Farm through SMS, apps such as Mkulima Bora)". Respondents then answered "yes" or "no" to the following m-services: Accessing farming (either livestock or agriculture) information, buying and selling agricultural or livestock products, and receiving important information or alerts on agriculture/livestock activities. We also asked respondents to provide the names of agricultural, livestock, or weather services apps accessed through their phones during the last growing season and included these summarized open-ended responses in the S.I[5].

[5]https://doi.org/10.6084/m9.figshare.12107166.v1

Table 2.1: Hypothesized relationships with m-services adoption

Independent Variables	Explanation
Personal smartphone	The rapid increase in mobile phone ownership in emerging and advanced economies–along with advances in internet speeds, global positioning systems, and cloud-based technologies–has led to ever-increasing quantities of digital data (World Bank, 2018a; GSMA, 2016). M-services content that contains multimedia formats or information accessible only in smartphone applications has correspondingly increased (World Bank, 2018a; Baumüller, 2018). Thus, personal smartphone ownership increases likelihood of m-services access because m-services are predominantly aimed for advanced devices.
Membership in farmer organizations	Increased knowledge and information exchange opportunities through government extension agencies and/or from farmer organizations increases m-services use. Kenyan farmers' membership in banana cooperatives has been shown to increase access to production information and innovation (Fischer and Qaim, 2012). Increased access and control over water resources explains technology adoption differences (Feder and Umali, 1993).
Farm size	Use of modern ICTs (e.g. mobile phones) among Indian farmers increases with farm size (Mittal and Mehar, 2016). Mittal et al. (2010) shows farmers with large plots better leverage information and communications than small farmers, e.g. benefiting most from obtaining information and accessing professional agronomic help due to fewer physical infrastructure constraints.
Household income	Willingness to bear financial risks of new technology performing poorly increases with income (Kebede, 1992).
Household and livestock assets	Wealthier households (those with greater quantities of livestock/household assets) may be better able to bear potential risks associated with technology adoption such as m-services (Kassie et al., 2013).
Age	Younger users tend to adopt new technologies more quickly than older users (Appelbaum, 1990). Willingness to invest in technologies with longer-term payoffs may decrease with age (Feder and Umali, 1993).
Education	Higher education levels influences an individuals' technology use (e.g. Piccoli et al. (2001)). Farmers' decisions to bear risks related to new technology adoption (Feder et al., 1985) and modern information sources such as mobile phones (Feder and Umali, 1993; Mittal and Mehar, 2016) positively correlates with education level.
County	We include County as a geographic variable to address potential variation in phone service quality that may affect m-service use.
Whether the respondent is a woman	Men and women perceive and adopt technology differently (Gefen and Straub, 1997). Lower literacy and education levels, household duties and heavy workloads, finances and lack of disposable income, social norms, all intersect to reduce women's mobile phone or internet use (Naidoo et al., 2008; Mwesige, 2004).

2.4 Data analysis

2.4.1 Description of variables

We selected three m-services as our dependent variables: farming and livestock information, buying and selling products, and alerts on agricultural or livestock activities. For brevity we refer to these as farming, buying and selling, and alerts, respectively. The following explanatory variables were taken from the household survey and grouped into the following categories. We note the theorized effect of the explanatory variables on the dependent variables in Table 2.1.

Personal Smartphone: We asked the farmer whether their mobile phone had access to the internet and the ability to download apps. If the farmer responded "yes" to both of those attributes, we classified them as owning a "smartphone" (i.e. a handset able to access the internet and download apps). We compared these to owners of non-smartphone phones which include basic and feature phones. "Basic" phones (*mulika mwizi* in Kiswahili) cannot download apps or access the internet. "Feature" phones are able to access the internet because they come pre-loaded with applications such as Facebook or Twitter but they do not have the ability to download apps. An example of a basic phone available for purchase in Kenya is Nokia 1110 whereas a feature phone pre-loaded with Facebook is Tecno T351. If respondents had more than one phone, we asked about their primary handset, i.e. the one used most often.

Membership in Farmer Organizations: We recorded membership in three types of farmer organizations (agricultural cooperatives, farmer groups, and Community Water Projects), and combined the responses to reflect the total number of farmer organization types (0-3) for which a farmer is a member.

Farm Size: Respondent provided the areas of their land under production and fallow, in acres, during the March-April-May 2018 growing season. We combined these two areas to get total farm size.

Livestock Assets: We asked the respondent to count the number of livestock owned by the household from a list of common animals including cattle, goats and sheep. We subsequently used a weighted formula to convert this livestock count to Tropical Livestock Units (TLU) (Jahnke, 1982).

Income: We asked the respondent to select household monthly income from a selection of ranges (none, 100-2,000 ksh, 2,001-6,000 ksh, 6,001-18,000 ksh, 18,001-36,000 ksh, 36,001-54,000 ksh, 54,001-72,000 ksh, or more than 72,000 ksh) from the following sources: casual labor, regular salary, small business, charcoal sales, horticulture, sale of forest products, livestock, remittances, rental income, pension, and savings group. We computed the income variables in multiple ways which produced similar results (see S.I.[6]) and ultimately selected the median value from those ranges. We summed the median values of the various incomes to estimate monthly income in Kenya shillings. We reduced the number of income categories by computing quartile incomes: 25% = 28,837 ksh, 50% 55,000 ksh, 75% = 86,575 ksh.

Assets: We used a simple index reflecting household ownership of a television, car, motorcycle, and/or computer (including tablets). We added the ownership values together and treated the sum as a continuous variable (0-4). However, because only four households owned all four assets, we combined those households with the households who owned three of the four assets (0-3).

Age: We asked for the respondent's year of birth and then calculated their age relative to the year 2018.

Education: Kenya follows the 8-4-4 educational system with eight years in primary school, four years in secondary, and four years of university or vocational training. We asked for the highest level of education completed by the respondent. The options were

[6]https://doi.org/10.6084/m9.figshare.12107166.v1

coded as follows: No formal education or some primary (level = 1; reference), Completed primary or some secondary (2), Completed secondary or some post-secondary (3), Completed post-secondary or vocational training (4).

County: Farmers were located in one of three counties: Laikipia, Meru, or Nyeri. We created dummy variables for Meru and Nyeri residents with Laikipia as the reference category.

Gender: Men were the reference category (coded as 0).

2.4.2 Logistic regression analysis to identify drivers of M-service adoption

We used a binary logistic Generalized Linear Model (GLM) to test the likelihood that a respondent adopts various m-services. We estimated the odds ratio (OR) for each dichotomous dependent variable: use or non-use of buying and selling information, information about agriculture and livestock, or important alerts on agriculture and livestock activities.

Because several farmers were part of the same water governance groups or farmer organizations, household-level data were not fully independent. Therefore, we also ran the regressions using a Generalized Linear Mixed Effects Model (GLMM) to account for possible overdispersion and clustering by requiring a group factor. Random effects modeled the correlation between the groups using community subsets called Water Resource User Associations (WRUAs) as the group factor. For more information regarding the treatment of WRUAs and grouping variables in the GLMM, see our Methodological Appendix (S.I.[7]). To test for the difference between the models with and without random effects, we used a nested ANOVA model comparison. We determined to proceed in using

[7] https://doi.org/10.6084/m9.figshare.12107166.v1

the GLM without random effects because the Aikake information criterion (AIC) values were lower for the GLMs compared to the GLMMs.

We cleaned data and developed the model specification using Python programming language (Python Software Foundation, 2019). The logistic regression analysis was completed using the lmer function in the lme4 library in R (Bates et al., 2015).

2.5 Results

2.5.1 Descriptive statistics

We begin by compiling the rates of phone ownership (smartphone, feature phone, basic phone) by gender to identify the extent to which lack of access to internet-enabled phones could underlie gender-based differences in the use of m-services in Table 2.3. Approximately 34 percent of respondents own a smartphone: a greater proportion of men own smartphones compared to women. Basic phones are owned by 56% of women compared to 48% of men. Considering that access to a smartphone is meaningful for m-services access, we investigated differences in age, education, and membership in farmer organizations between smartphone owners and non-owners, as shown in Table 2.4. The average education and number of farmer organizations is greater for smartphone owners compared to non-owners. The variance of those statistics is similar between smartphone owners and non-owners.

To help discern whether education, age, and membership in farmer organizations influences access to m-services we compare differences in these socioeconomic factors between men and women m-services users and non-users. As shown in Table 2.5, the average age and education level is lower for female m-services users than male. Similarly, women are on average members of fewer farmer organizations. The variances of those statistics

Table 2.2: Summary statistics of outcome variables and demographic characteristics of respondents. DV stands for Dependent Variable and IDV stands for Independent Variable.

Type	**Variable**	**Levels**	**n**	**%**
DV	Farming	No use	435	75.4
		Use	142	24.6
	Buying & selling	No use	445	77.1
		Use	132	22.9
	Alerts	No use	473	82.0
		Use	104	18.0
IDV	Personal smartphone	No	381	66.0
		Yes	196	34.0
	Membership in farmer organizations	None	100	17.3
		One farmer organization	263	45.6
		Two farmer organizations	172	29.8
		Three farmer organizations	42	7.3
	Membership in Community Water Project	No	125	21.7
		Yes	452	78.3
	Membership in informal farmer group	No	346	60.0
		Yes	231	40.0
	Membership in agricultural cooperative	No	500	86.7
		Yes	77	13.3
	Education (highest attained)	No formal education or some primary	105	18.2
		Completed primary of some secondary	240	41.6
		Completed secondary or some post-secondary	161	27.9
		Completed post-secondary or vocational training	71	12.3
	Income quartile[1]	<28,837 Ksh	144	25.0
		28,838 - 55,000 Ksh	135	23.4
		55,001 - 86,575 Ksh	143	24.8
		86,576 Ksh and above	155	26.9
	County	Laikipia	299	51.8
		Meru	150	26.0
		Nyeri	128	22.2
	Sex	Women	327	56.7
		Men	250	43.3

[1] Income is monthly and in Kenya shillings. We reduced the number of income categories by computing quartile incomes: 25% = 28,837 ksh, 50% 55,000 ksh, 75% = 86,575 ksh.

Table 2.3: Types of handset owned by respondent. Sample size for women $n = 327$ and men $n = 250$

Gender	Phone Type	No. of Respondents	Proportion
Women	Basic	184	56.27
	Feature	42	12.84
	Smartphone	101	30.89
Men	Basic	121	48.40
	Feature	34	13.60
	Smartphone	95	38.00

Table 2.4: Descriptive statistics of smartphone non-owners and owners membership in farmer organizations, age, and education

Variable	Statistic	Non-Owner	Owner
Farmer Organizations[1]	Mean	1.13	1.54
	σ [4]	0.823	0.780
Age[2]	Mean	52.6	47.8
	σ [4]	12.0	12.1
Education[3]	Mean	2.13	2.76
	σ [4]	0.839	0.918

[1] Farmer organizations reported in binary membership in farmer groups, CWP, and/or agricultural cooperative.
[2] Age reported in years.
[3] Education reported in highest level attained (see Description of variables for more information).
[4] Standard deviation

are generally similar between men and women. We use these differences between men and women to contextualize the importance of these socioeconomic factors in governing m-service use, despite them not always being significant in the logistic regression models.

Table 2.5: Summary statistics of m-services users and non-users by gender

	Alerts				Buying and selling				Farming			
	No use		Use		No use		Use		No use		Use	
	W[1]	M[2]	W	M	W	M	W	M	W	M	W	M
Age												
Mean	50.4	52.7	45.8	52.2	50.1	52.6	48.1	52.5	50.6	52.7	46.7	52.4
σ^3	12.22	11.98	10.39	13.74	12.57	12.40	9.67	12.38	12.42	12.30	10.18	12.63
Farmer organizations												
Mean	1.17	1.25	1.44	1.71	1.14	1.25	1.48	1.64	1.11	1.18	1.56	1.78
σ	0.829	0.815	0.769	0.803	0.835	0.838	0.731	0.753	0.800	0.810	0.828	0.736
Education [4]												
Mean	2.20	2.44	2.35	2.68	2.23	2.48	2.23	2.55	2.17	2.37	2.43	2.82
σ	0.939	0.933	0.635	0.811	0.938	0.907	0.745	0.926	0.941	0.887	0.714	0.893

[1] W, Women
[2] M, Men
[3] Standard deviation
[4] Education level is highest attained. Where no formal education or some primary (1); (reference), Completed primary or some secondary (2), Completed secondary or some post-secondary (3), Completed post-secondary or vocational training (4).

2.5.2 Logistic regression results

Smartphone ownership

Personal smartphone ownership, defined as owning a handset with access to internet and ability to download apps, increases the likelihood of m-services use in all three models as shown in Table 2.6. Smartphone owners are between 1.83 and 2.72 times as likely as non-smartphone owners to use m-services ($p < 0.05$ for buying and selling and $p < 0.001$ for farming, and alerts).

Farmer organizations

We find that membership in farmer organizations positively influenced m-service use. As shown in Table 2.6, respondents in farmer organizations are 1.64 to 2.06 times more likely than non-members to use m-services across the three types of farmer organizations, which include farmer cooperatives, informal farmer groups, and Community Water

Projects ($p < 0.001$). When disaggregating by farmer organization type in Table 2.7 our results show that members of informal farmer groups are between 1.82 and 2.87 times more likely to use farming (p $<$ 0.01), buying and selling ($p < 0.01$), and alerts ($p < 0.001$) m-services. However members of agricultural cooperatives are only more likely to use m-services for obtaining farming information ($p < 0.01$). Members of CWPs are only more likely to use m-services for accessing information about buying and selling ($p < 0.001$).

Wealth, gender and socio-economic factors

Household-level characteristics, such as household assets and livestock assets, and individual-level characteristics such as education are also associated with m-service use. After accounting for whether a respondent owns a smartphone, we do not find age or farm size to be significantly related with m-service use any model iterations ($p > 0.1$, Table 2.6 and Table 2.7). Income is significantly related to one m-service use in Tables 2.6 and 2.7 in the highest income quartile only. Other measures of wealth such as livestock assets (TLU), household assets, and the interaction between assets and TLU. Household and livestock assets are both significantly associated with likelihood of using m-services. Household assets are significantly associated with alerts m-services ($p <0.01$) while livestock assets are significantly associated with all three models in Table 2.7 and buying and selling and alerts m-services in Table 2.6. Although the coefficients on livestock assets and household assets were less than one, the interaction term was greater than one, pointing to a complex relationship between various forms of wealth and the adoption of m-services. As described in our Methodological Appendix (S.I.)[8], the results presented here are largely consistent with results of other iterations of the models, which controlled for interactions between wealth variables, including smartphone ownership.

[8]https://doi.org/10.6084/m9.figshare.12107166.v1

Additionally, in one iteration of the model, we replaced farm size as a proxy for income. We did not find any statistical significance for farm size or income.

We find some levels of educational attainment to be significantly correlated with m-service use in Table 2.6. Compared with the reference educational level, respondents who completed primary school or some secondary school were 3.16 and 4.55 times as likely to use farming and alerts m-services ($p < 0.01$). Similarly, respondents who completed secondary or some post-secondary school were 3.46 and 3.53 times more likely than the reference group to use farming and alerts m-services ($p < 0.01$ and $p < 0.05$, respectively). Respondents who completed post-secondary school or vocational training were 2.58 times more likely than the reference group to use farming m-services ($p < 0.05$). The results in Table 2.7 are similar for those education levels.

The coefficient on being a male respondent is positive in all iterations of the models (Tables 2.6, 2.7). In the models that do not distinguish between farmer organization types, Table 2.6, the association between gender and m-service use is significant for buying and selling m-services ($p < 0.01$) and alerts ($p < 0.001$). Men are 1.20 and 1.35 times more likely to use m-services for buying and selling and alerts.

We asked respondents who did not select use of m-services for accessing farming and livestock information to respond to a list that describes barriers preventing their use of m-services as shown in Figure 2.2. Compared to men, a greater proportion of women perceive barriers to mobile agriculture/livestock services across all of the categories. However, irrespective of gender, the foremost barriers limiting use of m-services are lack of awareness, lack of availability, and lack of understanding about how m-services work. Of these three categories, women are less aware and knowledgeable about how to use m-services compared to men.

Table 2.6: Predictors of use of farm/livestock information, to buy or sell agricultural produce or livestock, and alerts m-services

	Farming		**Buying & selling**		**Alerts**	
Predictors	*Odds Ratio*	*C.I.* [1]	*Odds Ratio*	*C.I.*	*Odds Ratio*	*C.I.*
Smartphone (Yes)	2.19[4]	1.39–3.45	1.83[2]	1.14–2.91	2.72[4]	1.62–4.55
Farmer Organizations	2.06[4]	1.56–2.72	1.64[4]	1.24–2.17	1.83[4]	1.33– 2.52
Farm Size	1.04	0.95–1.14	1.00	0.91–1.10	1.02	0.92–1.14
TLU	0.80	0.64–1.01	0.65[3]	0.50–0.85	0.51[4]	0.37–0.70
Income Quartile (Level=2)	1.10	0.58–2.08	1.65	0.86–3.16	1.75	0.84–3.63
Income Quartile (Level=3)	1.03	0.56–1.92	1.63	0.87–3.06	1.02	0.48–2.17
Income Quartile (Level=4)	1.48	0.79–2.74	1.70	0.89–3.25	2.30[2]	1.11–4.77
Assets	0.78	0.51–1.18	0.83	0.55–1.26	0.53[3]	0.33–0.83
Age	0.99	0.97–1.01	1.00	0.98–1.02	0.99	0.97–1.01
Education (Level=2)	3.16[3]	1.44–6.92	1.56	0.82–2.95	4.55[3]	1.69–12.29
Education (Level=3)	3.46[3]	1.54–7.77	1.13	0.56–2.26	3.53[2]	1.27–9.80
Education (Level=4)	2.58[2]	1.02–6.51	0.67	0.28–1.60	1.52	0.46–5.04
County (Meru)	1.05	0.63–1.75	1.63	0.99–2.68	0.90	0.50–1.65
County (Nyeri)	1.25	0.73–2.12	1.22	0.70–2.11	1.56	0.86–2.82
TLU*Assets	1.26	0.82–1.92	1.37	0.90–2.10	1.53	0.95–2.48
Sex (Male)	1.11	0.98–1.27	1.20[3]	1.05–1.38	1.35[4]	1.15–1.58
(Intercept)	0.05[4]	0.01–0.20	0.08[4]	0.02–0.30	0.05[4]	0.01– 0.26
Observations	577		577		577	

[1] Confidence Interval
[2] $p < 0.05$
[3] $p < 0.01$
[4] $p < 0.001$

Table 2.7: Model results which separates farmer organizations into farmer groups, agricultural cooperatives, and Community Water Projects (CWPs)

	Farming		**Buying & selling**		**Alerts**	
Predictors	*Odds Ratio*	*C.I.*	*Odds Ratio*	*C.I.*	*Odds Ratio*	*C.I.*
Smartphone (Yes)	2.25[4]	1.42–3.58	1.71[2]	1.06–2.75	2.69[4]	1.59–4.56
Farmer Group (Binary)	1.98[3]	1.29–3.04	1.82[3]	1.18–2.83	2.87[4]	1.73– 4.76
Ag Cooperative (Binary)	2.54[3]	1.44–4.47	0.78	0.41–1.47	0.97	0.48–1.96
Number of CWPs	1.38	0.94–2.03	1.95[4]	1.32–2.90	1.48	0.95–2.30
Farm Size	1.05	0.96–1.15	1.01	0.92–1.11	1.03	0.92–1.14
TLU	0.78[2]	0.62–0.99	0.66[3]	0.51–0.86	0.50[4]	0.37–0.69
Income Quartile (Level=2)	1.11	0.59–2.12	1.64	0.85–3.17	1.81	0.86–3.80
Income Quartile (Level=3)	1.10	0.59–2.05	1.59	0.84–3.02	1.01	0.47–2.16
Income Quartile (Level=4)	1.56	0.84–2.91	1.80	0.94–3.45	2.38[2]	1.13–4.99
Assets	0.79	0.52–1.20	0.82	0.54–1.24	0.53[3]	0.33–0.84
Age	0.99	0.98–1.01	1.00	0.98–1.02	0.99	0.97–1.01
Education (Level=2)	3.04[3]	1.39–6.68	1.55	0.82–2.96	4.80[3]	1.77–13.03
Education (Level=3)	3.37[3]	1.50–7.57	1.11	0.55–2.22	3.59[2]	1.29–10.01
Education (Level=4)	2.54[2]	1.00–6.42	0.67	0.28–1.61	1.48	0.44–4.98
County (Meru)	1.09	0.64–1.86	1.57	0.94–2.63	1.00	0.53–1.87
County (Nyeri)	1.44	0.83–2.48	1.25	0.71–2.19	1.83	0.99–3.38
TLU*Assets	1.27	0.83–1.94	1.47	0.96–2.27	1.62	1.00–2.65
Sex (Male)	1.11	0.97–1.26	1.19[2]	1.04–1.37	1.35[4]	1.15–1.58
(Intercept)	0.06[4]	0.01–0.24	0.06[4]	0.01–0.24	0.04[4]	0.01–0.22
Observations	577		577		577	

[1] Confidence Interval
[2] $p < 0.05$
[3] $p < 0.01$
[4] $p < 0.001$

2.6 Discussion

2.6.1 The role of smartphones in m-services use

We find smartphone ownership to be a significant factor in m-services use. However, smartphone ownership across our study site is far from ubiquitous with 31% ownership for women and 38% ownership for men (Table 2.3). GSMA (2016) estimates 226 million smartphone connections exist in Africa, approximately a quarter of all connections, and is reflected most strongly in established mobile markets including Kenya, Egypt, Nigeria, and South Africa. Thus, compared to the rates of smartphone adoption in east Africa (17% of the population in 2015) (GSMA, 2016), smartphone ownership is relatively high among survey respondents. The prevalence of smartphone ownership as a predictor of m-service use may be due to the nature of how m-services tend to be designed: technology developers generally design for smartphones rather than basic or feature phones (Wyche and Murphy, 2012; Cheney, 2018). Our results indicate that m-services, created by developers and designers focused on smartphone applications, have indeed reached the segments of the population who own smartphones. Considering that the majority of respondents own basic or feature phones (as shown in Table 2.3), m-services designed for smartphone-based applications may be failing low-income and basic phone-owning subscribers.

As suggested by several previous studies (e.g. Wyche and Steinfield (2016); Wyche and Murphy (2012); Wyche et al. (2019)), designing m-services with the needs of basic phone users in mind is a clear way to expand accessibility of m-services beyond smartphone users. Developers need not only to design for basic phones but also the constraints commonly experienced in developing countries. After relating barriers of basic phone use to lack of airtime credit, exhausted batteries, difficulty charging and lack of capital to upgrade to a smartphone, Wyche and Murphy (2012) provide an alternative design vision given the challenges faced by rural and often unconnected agriculturalists in Kenya. They suggest that mobile phone designers and developers assume off-grid use with unreliable electricity sources for agriculturalists in rural and peri-urban Kenya (Wyche and Murphy, 2012). Lastly, another consideration is to focus on USSD protocol and voice call services aimed for widely spread simple and cheap phones rather than power-intensive functionalities of smartphones which are not ubiquitously used (Wyche and Murphy, 2012). These intersections need to be addressed by ICT for development programs—such as those described in the USAID toolkit used for survey collection—for proper information dissemination.

2.6.2 Membership in farmer organizations increases likelihood of m-service use

Membership in informal farmer groups increased the likelihood of m-services use across all types of m-service use in comparison to formal farmer organizations (e.g. farmer cooperatives and CWPs). Unlike formal farmer cooperatives or the CWPs, informal farmer organizations are free to join. Yet, one benefit of paying to join an agricultural cooperative or CWP is access to information services and networks provided by these organizations. In other words, these types of formal organizations provide a 'club good' version of agricultural information (i.e., information exclusively available to members of an organization) (McNutt, 1999). Members gain exclusive access to and benefit from agricultural information at the price of paying a monthly membership fee to belong to these organizations.

Contrary to formal agricultural information, members of informal farming groups may make greater use of publicly-available agricultural information, which is a 'public good' (i.e., where an individual's use of information available to the public does not diminish others' use of the same information) (McNutt, 1999). Most m-services platforms do not require an explicit joining fee to access information and are therefore offering agriculture information as a publicly available good. Thus, we would expect that farmers in informal farming groups would have a greater incentive to use and seek out m-services agriculture information that they can freely access without having to pay the types of membership fees required of formal farmer organizations.

Farmers with membership in agricultural cooperatives or CWPs were significantly more likely to use only one type of m-service. Farmers in cooperatives already benefit from exclusive information and alerts that are embedded in their membership to a cooperative and the facilitated access to extension services. Thus, we would not expect these farmers to make substantive use of freely available m-services information. Rather than using the buying and selling function of m-services, farmers in formal cooperatives are likely already benefiting from their farmer cooperative membership with greater access to opportunities for marketing and buying and selling their products in higher-value markets (Narrod et al., 2009; Markelova and Mwangi, 2010). Similarly, farmers with CWP membership are likely to access agricultural information as an additional benefit to the primary benefit of obtaining water. In Kenyan CWPs, members interact on a weekly to monthly basis through labour activities for maintaining irrigation infrastructures and to attend meetings regarding irrigated water resources management. These activities would likely offer both informal and formal opportunities to obtain agricultural information. As with agricultural cooperatives, CWP membership likely provides access to exclusive agricultural information, and thereby reduces the incentive to access additional agricultural information via m-services.

By framing agricultural information access in the context of club and public goods, we can explain why farmers in informal farmer organizations are more likely to use m-services across all use types in comparison to those farmers belonging to agricultural cooperatives or CWPs. We conclude that membership in informal farmer groups is a strong predictor of m-services use due to their greater incentive to use publicly-available m-services compared to their agricultural cooperative and CWP counterparts, who inherently already have access to access these types of services via cooperative or CWP membership. We discuss endogeneity challenges related to farmer organizations and related explanatatory variables in the S.I[9].

2.6.3 Wealth indicators and education

The non-significant results for income and farm size, along with the ubiquity of mobile phones among respondents, point to the maturity of phone ownership in the study area relative to income. In earlier phases of mobile phone adoptions, higher income levels were associated with phone ownership, but over time, mobile phones have become accessible to even low-income households, as indicated by the 98% ownership of mobile phones among survey participants. However this ubiquity in phone ownership of all types does not represent a ubiquity of smartphone ownership, which is associated with higher incomes.

Although mobile phone ownership and use have been expanding across communities and throughout Kenya irrespective of wealth, smartphone ownership is the factor that separates the wealthy from the poor in m-service use. Moreover, smartphones open access to potentially the most comprehensive and/or useful m-services which are applications and/or internet-based platforms. Since Kenya is on the forefront of ICT use and mobile phone ownership compared to other countries in the region (see Figure 2.1), our results can serve as a model for future trajectories of development regarding use of mobile phones for climate and agriculture information. While these wealth indicators are non-significant across our models, we recognize that cost is still a major barrier to farmers using phones

[9] https://doi.org/10.6084/m9.figshare.12107166.v1

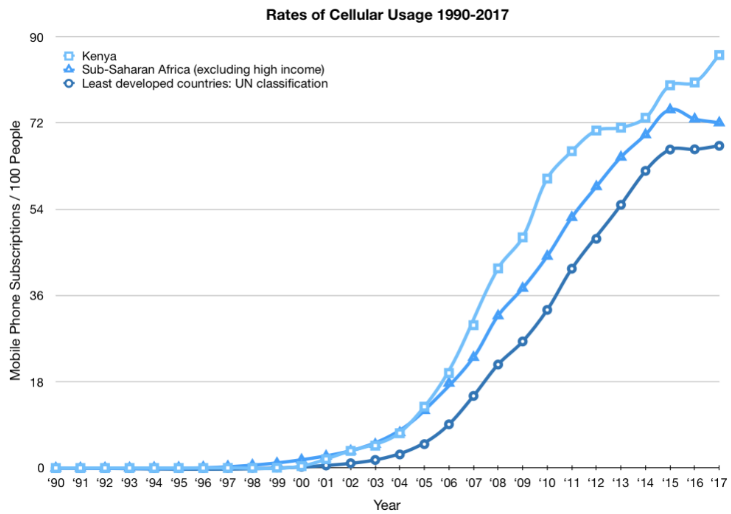

Figure 2.1: Growth rate of cellular usage as defined by mobile phone subscriptions per 100 people between 1990-2017. Mobile phone subscriptions in Kenya appear higher than mobile phone subscriptions of least developed countries according to UN Classification and Sub-Saharan African countries excluding high income. Data accessed on June 18, 2019. Data source: World Bank.

as much as they would like (see S.I.)[10]. We cannot conclude that access to smartphones means that farmers are using the internet. However given that purchasing airtime credit and/or data bundles is the chief reason why respondents do not use their phones (of any type) as much as they would like, we can infer that across all phone types, cost is still the greatest inhibitor of mobile phones and m-services use.

Educational attainment is an important factor in determining m-service use. The differences in average education level and age in m-service use for women, and to a lesser extent, men (Table 2.5) point at the importance of these socioeconomic factors in governing m-service use, despite them not always being significant in the models. While age does not significantly affect the likelihood of m-service use in our models, we do find that the average age of women using m-services is younger than those women who do not (Table 2.5), and the education level of those using m-services across both genders is

[10]https://doi.org/10.6084/m9.figshare.12107166.v1

higher than those who do not use m-services (Table 2.5).

2.6.4 Gendered barriers to m-service use

The previously described intersections between gender and wealth, smartphone ownership, and education affect m-service use in complex ways. At the household level, men's and women's roles on Kenyan farms are different, and therefore levels of agricultural decision-making may vary depending on gender roles when men are head of household. Braimok (2017) found in some cases that in the presence of a male head, female Kenyan dairy farmers did not perceive themselves as making or finalizing choices.

A variety of reasons govern these gendered barriers in use. Santosham and Lindsey (2015) concluded that cost is the greatest barrier to ownership and usage of mobile phones for women, due to their reduced financial independence compared to men. Although mobile phone use is nearly ubiquitous in our study site for both men and women, the increased costs associated with smartphone ownership and use present a potential barrier to women's ownership of smartphones, which may explain why women lag behind men in smartphone ownership (Table 2.3). Wyche and Olson (2018) found that mobile use and therefore m-service access among rural women remains limited due to technical literacy, mobile phone conditions, perceptions of the Internet, time required to learn how to use the Internet, and seasonal income fluctuations.

2.7 Implications for future m-services in Kenya

Membership in farmer organizations, the relative cost to own a smartphone, and smartphone usability are important considerations of future access to mobile agricultural services in Kenya. Although smartphone ownership is a driver of m-service use and smartphones are more expensive and difficult to use (Wyche and Steinfield, 2016), Karlsson

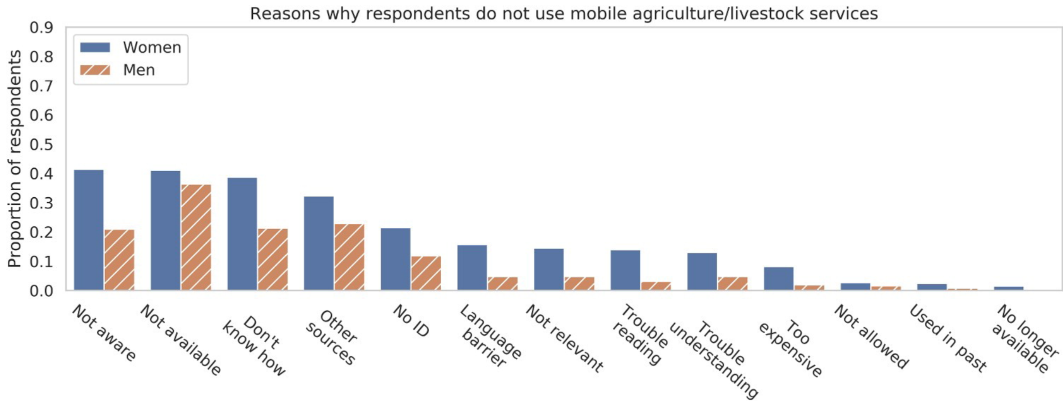

Figure 2.2: Bar plot showing barriers to m-service use by gender. We elicited specific reasons for farmers' non-use of mobile service in response to the question: What are the reasons why you don't access information about farming (either livestock or agriculture): I'm not aware of these services; These types of services are not available in my area/on my network; I don't know how to use these types of services; I get my information from other sources (e.g. my community); My phone has no internet; I do not have an ID or required documents; The content is not in a language I understand; The content isn't relevant to me; I have trouble reading the content; I have trouble understanding the content; They are too expensive; My family doesn't allow it; I've used it in the past but did not find it useful and/or did not like using it; I've used it in the past but it is no longer available.

et al. (2017) show that increasingly affordable smartphones are on the rise and commercially sound. African consumers can now purchase the affordable Mara X phone, which is designed in partnership with Google and fully manufactured in Rwanda[11] as well as Chinese designed smartphones that are expected to become more prevalent among African phone markets[12].

Beyond affordability, however, and from a supply point of view, m-services are not typically designed for low income, rural, and less educated groups, such as women, who use basic phones (Santosham and Lindsey, 2015; Wyche and Olson, 2018). Our quantitative findings corroborate previous qualitative studies and show that women were less likely to use m-services in comparison to men and that m-service users had higher levels of income and educational attainment. Increasing m-service use involves addressing both

[11]https://maraphones.com/blog/best-value-smartphone-africa/

[12]https://qz.com/africa/1374404/chinas-transsion-dominates-africas-phone-market-with-tecno-itel/

smartphone affordability and designing m-services for audiences with lower technological literacies.

Our study introduces additional novel findings that demonstrate linkages between adoption rates and the role of informal and formal farmer groups. Membership in increased numbers of farmer organization—formal or informal—was a significant predictor of m-service use. However, farmers not part of formal agriculture groups, such as agricultural cooperatives and water governance groups, were more likely to use m-services than farmers who were involved in informal farmer groups. Understanding that not all m-service users are embedded in formal agricultural groups is important for m-service design. Specifically, agricultural information shared amongst farmers in formal groups may be different than the m-services information accessed by farmers in informal groups. Overall, dissemination of meaningful and useful agro-meterological information via m-services requires consideration of the types of individuals that are currently accessing m-service information.

2.8 Conclusions

We investigated factors that influence farmer use of agriculture information from mobile phones, specifically m-services. While much of the literature in academic and government spheres stress the importance of age, location and education in defining the digital gender gap, our results highlight additional drivers of m-services use including participation in farmer organizations, higher-levels of education, and smartphone ownership. Each of these factors interacts with gender to highlight the disproportionate access of these services by women. Our results show that age and income are not significantly related to farmer use of mobile phone services; however smartphone ownership is a metric of individual wealth and assets that is significantly related to m-service use.

The prevalence of smartphone ownership across our models point to the trend that m-services are increasingly designed for advanced mobile phone features. Thus, we suggest that m-services providers design for the user, which is predominantly a basic or feature phone owner in the case of rural Kenyan farmers. Additionally, while rates of mobile phone ownership in Kenya are high, smartphone ownership is not as widespread. Although phones have been expanding through rural communities, smartphones are the item that separates the wealthier from the poorer. Smartphones can give farmers access to potentially the most useful of climate and agriculture related m-services. Mobile phones and other ICTs will continue to play important roles in managing risks and vulnerabilities associated with a changing climate (Eakin et al., 2015). Efforts in the climate adaptation policy space should address the general affordability of m-services and mobile phones, and better target under-served groups of users, particularly women and those not belonging to farmer organizations to improve efficient and targeted dissemination of agro-meteorological information services.

Chapter 3

Consequences of dryland maize planting decisions under increased seasonal rainfall variability

Abstract Shifts in rainfall frequency and intensity can lead to heavy crop loss in rain-fed agricultural systems. Small-scale farmers who plant with limited resources need to carefully select management strategies that are well suited for their environment. Farmers must choose between planting higher-yielding varieties that take longer to mature and lower-yielding varieties that can be harvested sooner. To better understand the interactions between rainfall variability, cultivar choice, and cropping success, we implement an ecohydrological model that accounts for variation in daily soil moisture and converts water stress to crop yield. We apply the model to growing conditions of dryland farmers in central Kenya, which is a drought-prone and semiarid region with spatially heterogeneous rainfall. To simulate stochastic daily rainfall, we derive parameters in 10-day increments from a 30+ year daily rainfall dataset. We use these properties to model the stochastic seasonal water availability for cultivars with different maturation lengths. In

agreement with past studies, our analysis shows that storms are becoming more intense and less frequent. We show that maize crops are prone to water deficit in the part of the growing season when crop water requirements are highest. Despite the potential for higher-yielding, late maturing varieties to improve total harvest, we find that early-maturing varieties that are drought-avoidant have the lowest likelihood of failure. In light of reduced rainfall totals, we show that the historical probability of crop failure was lowest in the past and is now increasing.

3.1 Introduction

In semiarid and arid regions of sub-Saharan Africa rising temperatures and shifting rainfall patterns are projected to negatively impact agricultural output (Downing et al., 1997; Slingo et al., 2005; Muller et al., 2011; Branca et al., 2011). Changes in rainfall associated with climate variability directly impact crop growth as storms are projected to become more intense with longer periods between rainfall events (Meehl et al., 2007; Donat et al., 2016; Harrison et al., 2019). The stochastic nature of rainfall during the growing season leaves crops susceptible to water stress during critical stages of development and can lead to crop failure (Sah et al., 2020; Salgado-Aguilar et al., 2020). Small rainfed farms cultivated by single families on plots less than 5 hectares represent the most prevalent form of agriculture in sub-Saharan Africa and are particularly vulnerable to climate variability and change (Samberg et al., 2016). Because of their dependence on rainfed agriculture (Dinar et al., 2008), smallholder livelihoods are susceptible to climate shocks that affect food prices (Ray et al., 2012), variability in production and supply (Lobell et al., 2011; Slingo et al., 2005), and farmer incomes (Reidsma et al., 2010).

Farmers make a variety of choices before and during the growing season that impact their agricultural production and thus food security and livelihood. Cultivar choice is one of the most critical choices a smallholder makes (Kalanda-Joshua et al., 2011). Because of the uncertainty associated with climate variability, farmer decision-making is becoming increasingly complex and uncertain at the expense of input use efficiency and profitability (Hansen et al., 2011; Waldman et al., 2019; Guido et al., 2020). Management options that were optimal for past or average climatic conditions may no longer be suited for increasingly common growing season weather. Additionally, traditional crop varieties may no longer be best suited for a farmer's environment, which has led to the development of hybrid and fast growing varieties (Smale and Jayne, 2010). Given ongoing changes in rainfall patterns, farmers need to select cultivars well suited for their local context that can lead to the greatest payoff in terms of yield while also minimizing the risk of crop failure. To date, however, no modeling exercise to understand the effect of stochastic rainfall variability on maize yields for various cultivars exists for dryland environments in sub-Saharan Africa.

To evaluate the impact of farmer decision-making and climate variability on agricultural outcomes, both field and modeling approaches have been employed (Bharwani et al., 2005; Ziervogel et al., 2005; Roudier et al., 2014; Vervoort et al., 2016; Wood et al., 2014; Choi et al., 2015). While field studies provide empirical evidence of environmental impacts on farmer outcomes, they can be limited to certain conditions, especially when panel data are absent (Patt et al., 2005; Hansen et al., 2011) and are difficult to extrapolate to scenarios where the climate is changing. Alternatively, crop models can be useful when field data are unavailable, but such models can also be limited in applicability and need to be carefully parameterized with special attention to how stochastic rainfall is modeled. However, given that daily rainfall observations with long temporal extents are generally unavailable for rainfed agricultural systems it is difficult to robustly estimate rainfall parameters. In rainfed contexts, variability in inter- and intra-annual rainfall is closely linked to variability in production. Crop models and agronomic studies in general focus on annual, seasonal or monthly rainfall totals (Barron et al., 2003) and do not provide a much-needed evaluation of interannual variability of within-season dynamics, which has important implications for crop yields (Recha et al., 2012). Dryland regions in particular necessitate careful modeling of rainfall patterns that are heterogeneous in space and time. An improved understanding of rainfall variability considers the temporal distribution of rainfall through analyses of the average amount of rain during rainfall events and the average length of time between successive events (Recha et al., 2012).

In order to develop a better way to study these systems, two concerns must be addressed. First a more accurate consideration of rainfall dynamics in semiarid environments is needed. Considering the stochastic nature of rainfall rather than the seasonal averages is important in these systems where the frequency and duration of rainfall lead to important consequences for vegetation response (Katul et al., 2007; Porporato et al., 2002). Localized convective storms arrive in pulses that beget nonlinear vegetation response (Katul et al., 2007; Baudena et al., 2007). Second, in addition to considerations of the hydroclimatic environment, the representation of vegetation needs to be specific to a crop of interest. While researchers have separately undertaken modeling exercises to understand the impact of climate variability on crop growth and the stochastic nature of rainfall on vegetation structure, there have been fewer efforts to link stochastic rainfall dynamics to the probability of crop failure for staple crops such as maize. Specifically, understanding the influence of cultivar choice on the success of a crop has not been considered.

This study is motivated by the need to better understand the coupled dynamics of water and rainfed agricultural systems in dryland regions occupied by smallholder farmers. In general, the relationship between soil moisture and maize yields is poorly resolved (Rigden et al., 2020). We address this need by presenting a model of stochastic seasonal soil water availability that evaluates the impacts of intra-seasonal rainfall variability on crop production in a smallholder agricultural system. This model is based on a previously explored stochastic soil water balance model (Laio et al., 2001b,a; Rodríguez-Iturbe et al., 2001; Porporato et al., 2001) that simulates the interactions between soil, plants, and climate. The point-based model is nonspatial and determines daily growing and harvest season values of runoff, interception, leakage and evapotranspiration for a given soil type and cultivar. Rainfall is represented as a marked Poisson process expressed as the mean depth of daily rainfall and the mean probability of storm arrival, which forces the model at the daily time step.

We apply this model to a study site in central Kenya that exhibits a high degree of rainfall variability. Using a long-term daily precipitation dataset and a characteristic soil type for the region we estimate climate and soil parameters for the model. We use the stochastic soil water balance model to determine yield outcomes and the probability of crop failure, which are a function of plant water deficit (dynamic water stress). We compare and evaluate our model results for various maize varieties with late, medium and early harvesting periods. We aim to answer the following questions for our study area:

1. What do historical records (40+ years) indicate about interannual rainfall trends?
2. How does intra-seasonal rainfall variability interact with the static and dynamic stress of a maize crop?
3. How do varying maize cultivars moderate the effect of climate variability on changes in yields and crop failure?
4. Has the average yield of maize production and likelihood of crop failure changed over an 80 year period?

The following paper is organized as follows: We first describe our study site and

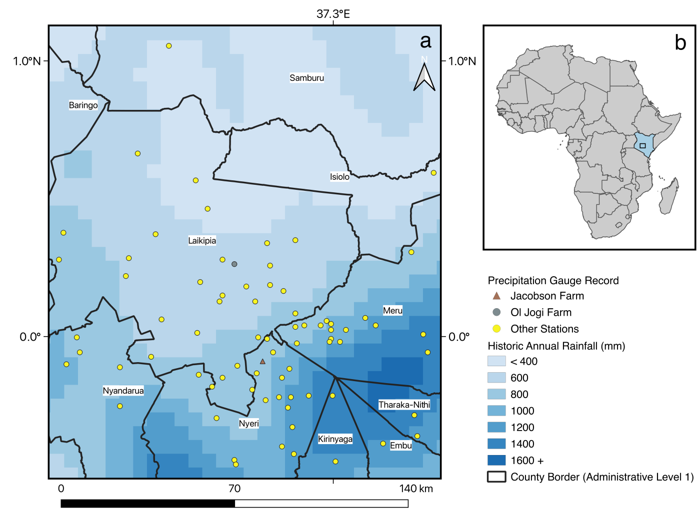

Figure 3.1: Maps showing: a) Inset of study site within central Kenya; b) Study site in Kenya noted by black box. Point data is from 80 stations. Jacobson Farm and Ol Jogi Farm are denoted with a triangle and circle icon, respectively. Shades of blue indicate historic annual rainfall based on enhanced Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) data that was created by blending 710 quality controlled station observations with the publicly available CHIRPS product averaged over 10-day periods in 1983-2016. Contours were delineated using the GeoClim Contour Tool using an interval of 200 mm.

the hydro-meteorological data necessary to apply the soil water balance model in the methods section. We then introduce our modeling framework and metrics for converting stress into yield. Our results demonstrate the impact of intra-seasonal rainfall variability on the seasonal water availability of maize varieties. We discuss these findings in the context of smallholder farmer decision-making and explore the implications of the results in the context of historical trends in rainfall and thus crop failure. We conclude with a discussion of model limitations and suggest additional research agendas appropriate for our proposed model.

3.2 Methods

3.2.1 Study site

We apply our model to the smallholder farming communities on the western slope of Mount Kenya, specifically the Laikipia plateau, in East Africa shown in Figure 4.1. Laikipia county is located on the western (leeward) side of Mount Kenya and is adjacent to Meru and Nyeri counties in central Kenya. The county comprises smallholder agriculturalists, growing urban areas, and wildlife conservatories that attract tourism. The presence of dryland agriculturalists along a heterogeneous rainfall gradient makes the area suitable for an analysis of rainfall variability and cropping outcomes. The main cropping season for maize is planting around day 100 (first week of April) and harvesting around day 300 (last week of October) (Ray et al., 2015).

Laikipia is a semiarid region prone to severe water deficits due to unreliable rainfall and high spatial and temporal variability. Rainfall in Kenya is characterized by a high coefficient of variability, which is common to semiarid environments (Herrero et al., 2010). Furthermore, Laikipia has a heterogeneous landscape and complex topography that results in a rainfall gradient. While the the foothills of Mount Kenya often receive between 800-900 mm of rainfall annually, the northern end of the county receives less than 500 mm annually (Wiesmann, 1998). The annual distribution of rainfall is bimodal with two rainy seasons: the long (roughly March through May) and short (roughly October through December) rains.

The Laikipia region of Kenya serves as an ideal study site to model the relationships between smallholder agriculture and climate variability for two reasons: (1) the tight couplings between food production and rainfall and (2) the prevalence of maize cultivation under various rainfall climatologies. In the region's drylands smallholder farmers face considerable challenges as rainfall arrives in pulses and in limited quantities for the majority of the year. Because the country has experienced a number of droughts in recent years, the Government of Kenya is especially interested in drought mitigation and increasing food security (Government of Kenya, 2010). Maize is an appropriate crop to study because it is grown under rainfed conditions and is an annual crop subject to both intermittent and terminal drought. Intermittent drought is caused by finite periods of inadequate water availability which does not necessarily result in crop failure whereas terminal drought is a progressive reduction in water availability that leads to crop failure before the end of the growing season (Neumann, 2008).

Soil types

We use the ISRIC Africa SoilGrids soil data base (Leenaars, 2014) to determine the soil textures found at depth 5-15 cm in our study site (Figure 4.1). The region has a heterogeneous mix of soil textures. The most prevalent soil textures, at the points of the rainfall gauges, are clay, clay loam and sandy clay loam. Soils in our catchment are geologically young soils derived from basaltic volcanic rock and are generally fertile but susceptible to erosion. The clay soils have high water storage capacities, which can be suitable for growing maize (Muchena and Gachene, 1988). In Table 3.1 we show the corresponding values for the soil matrix potential at the hygroscopic point Ψ_{s_h}, and at field capacity $\Psi_{s_{fc}}$, the porosity n, and the saturated hydraulic conductivity K_s according to the values found in Clapp and Hornberger (1978).

Interannual trends in long-term rainfall records

To assess whether interannual trends are changing across Laikipia, we use long-term records of daily rainfall data for stations across the study site provided by the Centre for Training and Integrated Research in Arid and Semi-Arid Lands Development (CETRAD) in Nanyuki, Kenya. The gauges, shown in Figure 4.1, have record lengths between 7 to

79 years. For computing statistics in Table 3.4, we used stations with records of 40+ years. We considered temporal trends in the two parameters: the average depth per rain event (α) and the average rain event frequency per day (λ) during the two rainy seasons: long (March-May) and short (October-December) rains. We used these daily records to calculate mean rain depth per event (α), and rain event frequency (λ), values for the regional trend analysis shown in Table 3.4, which is different than the calculation of 10-day estimates of rain depth and rain event frequency described in section 3.2.2 and used for the model parameters.

To analyze interannual trends in the total seasonal rainfall, rain depth per event and rain event frequency parameters for the two seasons, we used a modified Mann-Kendall statistical test and the Theil-Sen estimator. We use a variance corrected Mann-Kendall test proposed by Yue and Wang (2004), which is appropriate for rainfall data and calculates an effective sample size using the lag-1 autocorrelation coefficient. We used the modified-mk package in R (Patakamuri and O'Brien, 2020).

Maize varieties and yields

In addition to being a function of environmental conditions and management decisions, total yields also depend on the maize variety. To define maximum potential yields for our simulations, we use empirical yield potentials for maize varieties typically grown by smallholder farmers in Laikipia. These data were sourced from Kenya Seed Company (see Figure 3.3). The gap between yields reported by seed companies which were grown under controlled conditions and on small-scale farmers' fields is likely greatest in sub-Saharan Africa compared to other regions (Setimela et al., 2017). Evidence from Zambia shows that yields attained by seed companies in optimal settings and the actual realized yields by small-scale farmers can be vastly different (Blekking et al., 2021). Thus, the values of yields used in our model should be considered as the "potential" yield attained

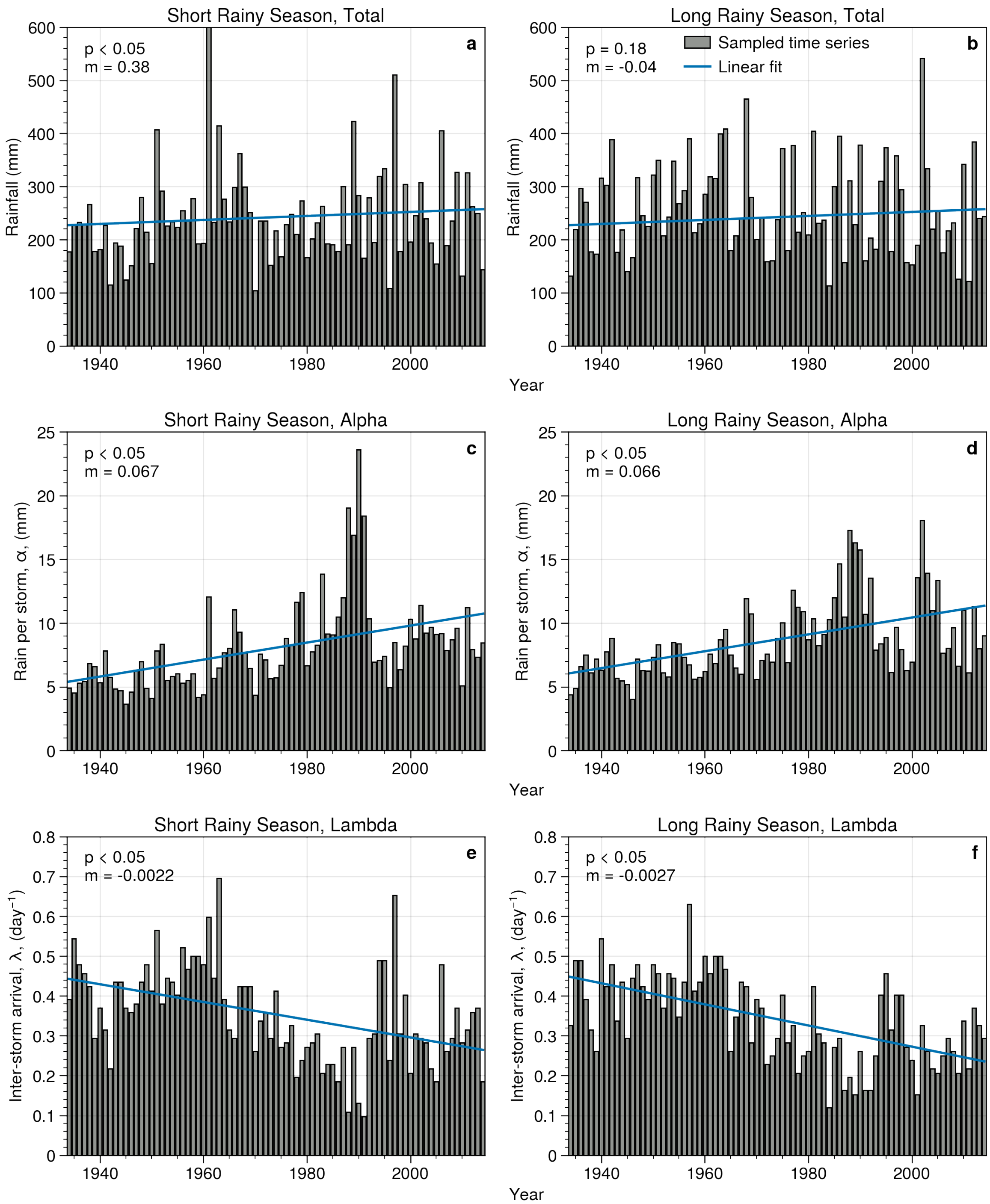

Figure 3.2: Time series for the Jacobson Farm station which has a 79-year record length. Significant trends (p <0.05) are shown in plots a, c, d, e, f per the modified Mann-Kendall test.

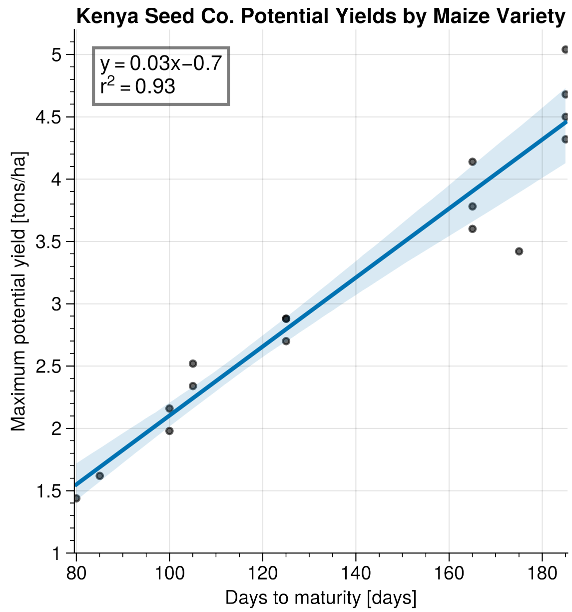

Figure 3.3: Linear regression with 95% confidence interval. Data are maize varieties sold by Kenya Seed Company. We removed two varieties from the dataset (PH1 and PH4) because the days to maturity information from Kenya Seed contradicted what is known about these four-month varieties. Data retrieved August 19 2019 from https://web.archive.org/web/20190216031348/http://kenyaseed.com/gallery/maize/.

by farmers rather than what is realized in-field. We discuss the uncertainty between these measures in A.1. We use the linear regression trend to set maximum potential yields for a range of maize varieties with maturity periods between 80 and 180 days, which is further discussed in section 3.2.3. For each maize variety, we calculate a maximum potential yield, Y_{max}, which is used in Eq. 3.17.

3.2.2 Model description

Hydrological water balance

To characterize the field-scale water balance we use a simple single-layer model of soil moisture at a point. The overall water balance is given by:

Table 3.1: Parameters associated with three prominent soil textures found within the study site.

Soil type	Ψ_s (MPa)^b	b^b	K_s (cm/d)^b	n^b	s_h	s_{fc}
Clay	-3.97×10^{-3}	11.4	11.1	0.482	0.503	0.830
Clay loam	-6.17×10^{-3}	8.52	21.2	0.476	0.420	0.821
Sandy clay loam	-2.93×10^{-3}	7.12	54.4	0.420	0.319	0.711

[a]We calculated values of s_h and s_{fc} assuming a soil water potential $\Psi_h = -10.0$ MPa and $\Psi_{sfc} = -0.03$ MPa (Laio et al., 2001b).
[b]Clapp and Hornberger (1978).

$$nZ_r \frac{\mathrm{d}s(t)}{\mathrm{d}t} = R(t) - I(t) - ET(s(t)) - L(s(t)) - Q(s(t)), \tag{3.1}$$

where n is porosity [-], Z_r is the rooting depth [mm], $s(t)$ is the saturation, or relative soil moisture content ($0 \leq s(t) \leq 1$), $R(t)$ is the rainfall rate [mm day^{-1}], $I(t)$ is the amount of rainfall lost through canopy interception [mm day^{-1}], $E(s(t))$ is the rate of evapotranspiration [mm day^{-1}], $L(s(t))$ is the rate of leakage [mm day^{-1}], and $Q(s(t), t)$ is the runoff rate [mm day^{-1}]. This is analogous to Eq. 1 from Rodriguez-Iturbe, et al. (2001).

Rainfall How rainfall varies within seasons and between seasons is highly important to dryland agriculturalists, and we capture both of these aspects of rainfall variability in our model. First, we wanted to capture the within seasonal variability of rainfall, which we model using two parameters: the average depth per rain event, α and the rain event frequency, λ. Rainfall is treated as a time-varying marked Poisson process. In each 10-day increment i, rainfall occurs with a probability, λ_i [day^{-1}], and the depth of rainfall events are drawn from an exponential distribution with a mean value, α_i [mm]. Values of λ and α used to model daily rainfall are estimated over 10-day fixed window intervals from a long-term rainfall record of daily station rainfall data, which was described previously in section 3.2.1. To capture the seasonal variation in rainfall parameters, we selected a

Table 3.2: Characteristics of rain gauges used in model simulations.[a]

Site	Latitude	Longitude	Mean Annual Rainfall, mm	Altitude, m.a.s.l.	Start Year	End Year
Jacobson Farm	0.04	37.04	735	1875	1934	2014
Ol Jogi Farm	0.31	36.94	538	1741	1967	1999

[a]Start and end years are the years when the rainfall records began and ended for each rain gauge.

10-day fixed window instead of monthly intervals to estimate α and λ, which are then used to simulate daily rainfall during the season. The choice of a 10-day interval allowed rainfall parameters to vary within each month, which is important for capturing the onset and end of the rainy seasons in this region. However, even with multiple decades of data (e.g. 31 years of data at the Ol Jogi station), there are not enough observations of rainfall to justify daily parameter estimates.

Second, we wanted to model the interannual variability of total seasonal amounts in a way that most closely represented observations from our rain gauge record. We found that using a fixed rainfall climatology across seasons led to underestimations of the variance in seasonal rainfall from year-to-year and did not accurately represent the observed coefficient of variation in seasonal rainfall for our rain gauge locations of interest. As described earlier, our intraseasonal estimates of λ and α are average values derived from multi-year averages of the rainfall process for each 10-day interval. However it is likely that inherent variability in these parameters from year-to-year leads to increased variance in seasonal totals compared to what we see when using the same average values for each season. To account for this additional variability in the rainfall process between seasons, we modify all of the 10-day λ increments for each season of our simulation using multiplicative Gaussian noise. Specifically, for each season, we multiply the original λ_i for each 10-day time interval by the same seasonal scale factor, which is drawn randomly for each season from a normal distribution with a mean of 1 and a standard deviation of 0.35. The scale factor of 0.35 was selected in order to match the coefficient of variation and seasonal totals from the Ol Jogi and Jacobson Farm stations. We recognize that there are many possible mechanisms that could increase the variability of seasonal rainfall totals and that some of these mechanisms would be tied to the daily rainfall process itself in ways that would require modification of the time-varying poisson model we have proposed. However, the addition of the random multiplicative noise offers a simple approach for capturing interseasonal variability in rainfall without making additional and more complex assumptions regarding the probabilistic nature of the daily rainfall process.

Interception and Evapotranspiration In maize systems, rainfall interception by the canopy is proportional to canopy Leaf Area Index (LAI) and will therefore be highest during tasseling (reproductive) and maturity stages (Zheng et al., 2018). We included a simple estimate of interception $I(t)$ such that

$$I(t) = LAI * I_{\mathrm{e}} \tag{3.2}$$

where the crop LAI, defined in equation 3.5, is transformed linearly by an interception efficiency term I_{e} which converts the units of LAI to canopy interception [mm day^{-1}]. Zheng et al. (2018) found that canopy interception loss accounted for almost 50 percent of rainfall events less than 5 mm in a semiarid region in China. In our model, the maximum interception value is the maximum value of LAI, which is 3 mm.

Evapotranspiration depends on both the soil moisture, s, and the time into the growing season, t. We separate evapotranspiration into its two components, soil evaporation $E(s,t)$ and plant transpiration $T(s,t)$ such that

$$ET(s,t) = E(s,t) + T(s,t). \tag{3.3}$$

Soil evaporation depends on both the amount of soil moisture and the extent of the crop canopy, which intercepts radiation and reduces energy available for soil evaporation. We define evaporation as

$$E(s,t) = \begin{cases} 0 & 0 \le s < s_{\mathrm{h}} \\ \left(\frac{s-s_{\mathrm{h}}}{1-s_{\mathrm{h}}}\right)^{q_{\mathrm{e}}} ET_{\max} \mathrm{e}^{-0.5 \cdot LAI(t)} & s_{\mathrm{h}} \le s \le 1 \end{cases}, \tag{3.4}$$

where q_{e} represents the non-linear rate at which soil evaporation declines as soil moisture drops below saturation, LAI denotes the crop leaf area index [mm^2 mm^{-2}], a measure of canopy density, and s_{h} is the soil moisture at the hygroscopic point. ET_{max} is both the maximum PET and maximum bare soil evaporation. $LAI(t)$ varies depending on the time into the growing season, t, and is calculated from the maximum leaf area index of a given crop, $LAI_{c,max}$ and the crop coefficient, $K_c(t)$, as follows:

$$LAI(t) = LAI_{\mathrm{c,max}} \frac{K_c(t)}{K_{c,\mathrm{max}}} \tag{3.5}$$

We use a convex function of soil moisture for evaporation because with a 400 mm depth uniform layer we cannot easily resolve the problem of different evaporation curves for a two layer model. Rather, we consider that evaporation drops off relatively fast below saturation and thus we might be under estimating evaporation in certain cases such as in large rainfall events early in the season or when the soil is very wet, which is a method applied in previous studies (Caylor et al., 2005; Porporato et al., 2003). Many bucket models neglect early season evaporation all together, and our model is an improvement to this. Future research could try to make a bucket model with Stage 1 and Stage 2 evaporation but we have found that this is not usually done in models but rather is included in empirical studies of deep soil moisture.

Seasonal variation in the Crop coefficient The growing season is divided into four stages of crop phenology, through which the crop coefficient varies, peaking in mid-season after the reproductive stage. The crop coefficient is determined as

$$K_c(t) = \begin{cases} K_{c,\mathrm{ini}} & t \leq f_\mathrm{i} \\ \frac{K_{c,\mathrm{max}}-K_{c,\mathrm{ini}}}{f_\mathrm{d}-f_\mathrm{i}}(t-f_\mathrm{i}) + K_{c,\mathrm{ini}} & f_\mathrm{i} < t \leq f_\mathrm{d} \\ K_{c,\mathrm{max}} & f_\mathrm{d} < t \leq f_\mathrm{ms} \\ \frac{K_{c,\mathrm{eos}}-K_{c,\mathrm{max}}}{f_\mathrm{ls}-f_\mathrm{ms}}(t-f_\mathrm{ms}) + K_{c,\mathrm{max}} & f_\mathrm{ms} < t < f_\mathrm{ls} \\ K_{c,\mathrm{eos}} & t = f_\mathrm{ls} \end{cases} \tag{3.6}$$

where f_i, f_d, f_ms, and f_ls denote the fraction of the growing season in days from the beginning of the growing season to the vegetative period (initial), reproductive period (development), maturity period (mid-season), and end of season (late season), respectively. The values of K_c during vegetative and maturity periods are constant functions of time and the values of K_c during reproductive and senescence stages are linearly interpolated between the values for the start and end of each period. We use crop coefficient (K_c) values based on those listed in FAO guidelines for 180-day maize (grain) in high altitude East Africa, which are defined as 0.3, 0.3, 1.2, 1.2, and 0.6 and correspond to 0%, 16%, 44%, 76% and 100% of the growing season (Allen et al., 1998). We selected this metric for K_c because it is widely used in the Crop Water Requirement Satisfaction Index (WRSI) (Senay, 2004) when locally appropriate values are absent.

The rate of plant transpiration, $T(s)$, is given as

$$T(s) = \begin{cases} 0 & s < s_\mathrm{w} \\ \frac{s-s_\mathrm{w}}{s^*-s_\mathrm{w}} K_c T_\mathrm{max} & s_\mathrm{w} \leq s < s^* \\ K_c T_\mathrm{max} & s^* \leq s \leq 1 \end{cases} \tag{3.7}$$

where T_max is the maximum transpiration rate of the plant, s_w is the soil moisture at the

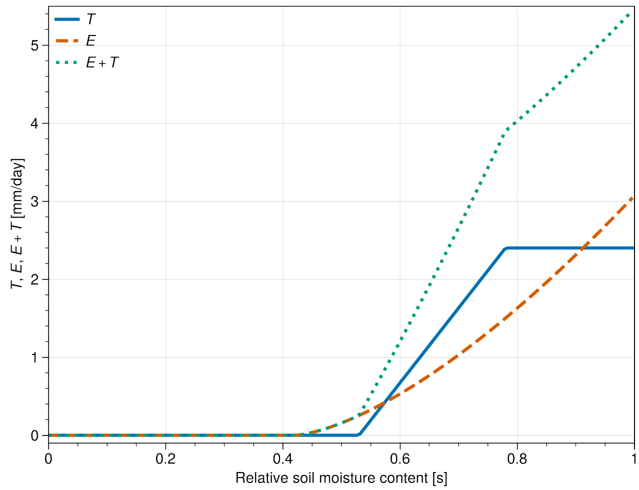

Figure 3.4: Evaporation, transpiration, and evapotranspiration as functions of relative soil moisture where LAI is 1.5 and s_{h} is 0.42. Other climate and crop parameters are those listed in Table 3.3.

wilting point, and, s^* is the soil moisture at the stress point.

Having solved for the components of Eq. 3.3, we can use Eqs. 3.4 and 3.7 express $ET(s)$ as follows:

$$ET(s) = \begin{cases} 0 & 0 \leq s < s_{\mathrm{h}} \\ \left(\frac{s-s_{\mathrm{h}}}{1-s_{\mathrm{h}}}\right)^{q_{\mathrm{e}}} ET_{\max} \mathrm{e}^{-0.5 \cdot LAI} & s_{\mathrm{h}} \leq s < s_{\mathrm{w}} \\ \left(\frac{s-s_{\mathrm{h}}}{1-s_{\mathrm{h}}}\right)^{q_{\mathrm{e}}} ET_{\max} \mathrm{e}^{-0.5 \cdot LAI} + \frac{s-s_{\mathrm{w}}}{s^*-s_{\mathrm{w}}} K_c T_{\max} & s_{\mathrm{w}} \leq s < s^* \\ \left(\frac{s-s_{\mathrm{h}}}{1-s_{\mathrm{h}}}\right)^{q_{\mathrm{e}}} ET_{\max} \mathrm{e}^{-0.5 \cdot LAI} + K_c T_{\max} & s^* \leq s \leq 1 \end{cases} \tag{3.8}$$

Figure 3.4 shows the functional form of evaporation and transpiration as a function of relative soil moisture content for a clay loam soil in which s_{h} is 0.42, s_{w} is 0.53, and s^* is 0.78.

Leakage Whenever daily soil moisture exceeds the soil field capacity, s_{fc}, we calculate leakage of water out of the plant root zone. The instantaneous rate of leakage is determined by the hydraulic conductivity, K(s) [mm/day], which is given by

$$L(s) = K(s) = \frac{K_s}{\mathrm{e}^{\beta(1-s_{\mathrm{fc}})} - 1}(\mathrm{e}^{\beta(s-s_{\mathrm{fc}})} - 1), \tag{3.9}$$

where K_s is the saturated hydraulic conductivity [mm/day] and β is a soil-specific parameter that governs the shape of the relationship between saturation and hydraulic conductivity; that is, $\beta = 2b + 4$, where b is a coefficient governing the power-law form of the soil-water retention curve (Laio et al., 2001b).

$$\Psi_s = \overline{\Psi}_s \, s^{-b} \tag{3.10}$$

where Ψ_s is the soil matrix potential at a given value of soil moisture and $\overline{\Psi}_s$ is the geometric mean of the soil matrix potential values in the curve. Values of b are based on soil texture and are taken from the empirically-determined coefficients presented in Clapp and Hornberger (1978).

Following an input of rainfall such that $s_0 > s_{\mathrm{fc}}$, Eq. 3.1 becomes an expression of soil moisture decay. In the absence of evaporative losses, we can simplify leakage to

$$L(s) = -nZ_{\mathrm{r}}\frac{\mathrm{d}s}{\mathrm{d}t} \tag{3.11}$$

Then we use Eqs. 3.9 and 3.11 to solve for the initial condition, s_0, which yields the analytical solution for total daily leakage when $s > s_{\mathrm{fc}}$:

$$L = \frac{nZ_{\mathrm{r}}}{\beta}\ln\left[\mathrm{e}^{\beta(s_0-s_{\mathrm{fc}})} - \mathrm{e}^{-m\beta}(\mathrm{e}^{\beta(s_0-s_{\mathrm{fc}})} - 1)\right] \tag{3.12}$$

where

$$m = \frac{K_s}{nZ_{\mathrm{r}}\left(\mathrm{e}^{\beta(1-s_{\mathrm{fc}})} - 1\right)} \tag{3.13}$$

Runoff Our model only considers saturation excess overland flow, so that when the balance of daily rainfall, evaporation, and leakage leads to an excess of soil saturation, the excess is converted to surface runoff. Thus, we can write

$$Q(s) = \begin{cases} 0 & 0 \leq s \leq 1 \\ (s-1)nZ_{\mathrm{r}} & s > 1 \end{cases} \tag{3.14}$$

Plant water stress

Water stress during the growing season affects the physiology of plants and is crucial in determining the success of a crop. Soil moisture excursions below the wilting point, s_w, and the point at which transpiration is reduced, s^*, can lead to reduced biomass production and eventual crop failure. By defining the duration and frequency of soil moisture deficits probabilistically, we gain insight into the yield reductions from intense water stress such as drought and the likelihood of crop failure. The mathematical derivations for these dynamics are described in previous work (e.g. Rodríguez-Iturbe and Porporato (2007).)

First we calculate static water stress: a measure of the mean vegetation water stress incurred during a given excursion below s^* (Porporato et al., 2001).

$$\zeta(t) = \left[\frac{s^* - s(t)}{s^* - s_w}\right]^{q_{\mathrm{stress}}}, \text{ for } s_w \leq s(t) \leq s^* \tag{3.15}$$

where $\zeta(t)$ is the static water stress and q_{stress} represents the crop's sensitivity to the magnitude of excursions below the stress point.

Although the static stress describes the mean water deficit relative to s^* and s_w, we are also interested in the duration and frequency of water deficit. Following the work of (Rodríguez-Iturbe and Porporato, 2007, p. 67), we also define two other random variables: the average length of time in days in which soil moisture is below the threshold, T_ζ, and the number of times the threshold is crossed during a season, n_ζ. Next, we calculate dynamic water stress, θ, a measure of the crop's total water stress during the growing season that characterizes the duration and frequency of exposure to water stress:

$$\theta = \begin{cases} \left(\frac{\bar{\zeta}\,\overline{T_{s*}}}{kLGP}\right)^{n_{s^*}^{-r}} & , \bar{\zeta}\,\overline{T_{s*}} < kLGP \\ 1 & , \bar{\zeta}\,\overline{T_{s*}} \geq kLGP \end{cases} \tag{3.16}$$

where $\overline{\zeta}$ is the average static water stress incurred over the growing season; $\overline{T}_{s*}$ is the average duration of excursion; n_s is the number of occurrences of these excursions during the growing season; LGP is the length of the growing period in days; k is the portion of the season that stress can occur before the crop fails; and r is a normalization parameter for the number of excursion below the stress point. We use the values for r and k noted in Table 3.3. The r parameter is defined as in Porporato et al. (2001), and the k parameter is selected in order to approximate a characteristic rainfall yield relationship such as in Guan et al. (2017). Lastly, we calculate seasonal crop yields, Y, as a function of dynamic water stress:

$$Y = Y_{max}(1 - \theta) \tag{3.17}$$

where the maximum yield per unit area (hectare) is Y_{max}. This value of Y_{max} is specific to the LGP of the crop as estimated by the linear regression in Figure 3.3. The maximum potential yield for a variety with a 180 day LGP is 4.26 metric tons per hectare. We conducted a sensitivity analysis to test different values of r, k, q_{stress}, and q_e on yields in

Table 3.3: Model parameters and their sources.

Type	Parameter	Value
Climate parameters	LGP	180 days
Crop parameters	I_{e}	1 (dim)
	Z_{r}	400 mm[a]
	T_{max}	4.0 mm day^{-1}
	ET_{max}	6.5 mm day^{-1}[b]
	$K_{c,\mathrm{max}}$	1.2 (dim)[c]
	LAI_{c}	3.0 mm^2 mm^{-2}[d]
	q_{e}	1.5 (dim)
	q_{stress}	2 (dim)
	r	0.5 (dim)[e]
	k	0.25 (dim)
	Y_{max}	4.26 t ha^{-1}
Soil parameters (clay loam)	s_w	0.53 (dim)[f]
	s^*	0.78 (dim)[f]
Simulation parameters	N_{sim}	10,000 seasons
	$N_{\mathrm{burn\_sim}}$	1,000 seasons
	T_{burn}	60 days

[a]Nyakudya and Stroosnijder (2014)
[b]Barron et al. (2003)
[c]Allen et al. (1998)
[d]Williams and Albertson (2004)
[e]Porporato et al. (2001)
[f]Clapp and Hornberger (1978)

Figure A.1 and discuss the need for improved measurements of intra-seaonal crop stress in A.2.

3.2.3 Model implementation

In order to conserve the water balance, we implement the model piecewise. First, we calculate saturation excess, $Q(t)$, which is caused by an input of rainfall that leads to the soil moisture being above saturation (i.e. $s > 1$). Therefore, any rainfall that causes s to be greater than 1 is instantaneously routed as runoff, and thus there is an explicit upper bound of s which is 1 as shown in equation 3.8. We then calculate both leakage and evapotranspiration using the same s value and within the same time step. Lastly, once we have values of $L(t)$ and $ET(t)$, we update the water balance and the soil moisture.

We run the model for 10,000 simulations using the following initial conditions and parameters. We use the Ol Jogi Farm rainfall climatology (10-day rainfall parameters estimated from daily rainfall records between 1967-1998), a clay loam soil texture, and 180-day maize variety, which are typical values for the study site. We selected a planting date of March 1 (Julian day 60) for model calibration and simulations. Farmers vary in their planting practice but generally plant at the start of the long rains after approximately 20 mm of rainfall has fallen. We selected March 1 because it was the most frequently planted week for the long rains among surveyed farmers in our study site. To determine the initial condition for our simulations, we first run the model for sixty days (T_{burn}) before the planting date and run 1,000 simulations ($N_{\text{burn\_sim}}$) to get an average value of s for the first day of the season. Then we use this average value as the initial condition for each of our subsequent simulations of the growing season.

Impacts of maize variety on farming outcomes

We continue with the model parameters previously described to compare yield outcomes for different maize varieties (e.g. Ol Jogi rainfall climatology, clay loam soil, planting day 60 etc.). We run the simulations for three categories of varieties: early, medium, and late maturing. The simulations have the same precipitation forcing for all the varieties, and the only lever that is being changed for this experiment is the length of the growing period (LGP). The maize varieties have different times to maturity that vary in five-day increments between 80 and 180 days: early maturing between 80 and 110 days, medium maturing between 115 and 145 days, and late maturing between 150 to 180 days (all inclusive). We designate 21 different maize varieties to represent all of the five-day increments between 80 and 180 days. These 21 varieties are each simulated

500 times, resulting in 10,500 total simulations.

Impacts of rainfall climatologies on farming outcomes

We use two rainfall gauges to determine the 10-day average rain depth per event (α) and rain event frequency (λ) used as model parameters parameters, which are described in Table 3.2. First, we use the long-term average Ol Jogi Farm rainfall climatology for running the model as this location is climatologically representative of semiarid small-scale producers in the region who depend on rainfall. We use this station for the model simulation results in sections 3.3.1 through 3.3.3.

We use the longest station gauge record (79 years), Jacobson Farm rainfall climatology, to analyze long-term temporal changes in crop production and crop failure in section 3.3.5. We define three eras of crop production: two extreme conditions (1930s rainfall climatology and 2010s rainfall climatology) and the average rainfall conditions. We extract 10-day rain depth per event (α) and rain event frequency (λ) values as follows. To represent average annual change in either parameter, we use the average 10-day α and λ values, which considers the entire rainfall record and corresponds to the climate in the middle of the record. We adjust the long-term average values for each 10-day increment in order to obtain historical (1930s) or present (2010s) values of the parameters 10-day increment. To estimate the past, "1930s", 10-day rain depth per event (α) and rain event frequency (λ) values we simply apply the trend line to each of the individual average 10-day increment. For example, we calculate the 10-day rain depth per event (α) for the 1930s climatology by taking the average 10-day alpha and subtracting it by the slope of the trend line multiplied by 40 years. For the present "2010s" values we add the slope of the trend line multiplied by 40 years. We then use these three sets of average rain depth per event (α) and rain event frequency (λ) values to run the simulations for the three subsetted periods. We run 100,000 simulations for each period with the parameters used in Table 3.3. Therefore, the only source of variability in this counterfactual analysis of crop failure and yields is the rainfall generator.

3.3 Results

3.3.1 Time-varying impacts of seasonal water availability using the Ol Jogi rainfall climatology

Figure 3.5 (bottom panel) shows the average soil moisture content for 10,000 growing seasons of a 180 day variety planted on day 60 (approx. March 1) using the Ol Jogi rainfall climatology. The 10-day rain depth per event (α) and rain event frequency (λ) parameters used to generate rainfall are variable over the 6 month period. The rain event frequency, λ, increases more than three-fold during the first 80 days of the season while the average depth per storm, α, stays relatively constant throughout the 180 day growing season with values between 8-13 mm of rain (Fig. 3.5, top). On dekads 15 and 16 (approx. 80-90 days into the growing season), the rain event frequency, λ, drops and levels off indicating the cessation of the long rains season. During this period of rainfall variability, the crop coefficient follows a step-wise function in which water requirements are low before steadily increasing when the crop begins to develop (Fig. 3.5, middle plot).

For the first 80 days of crop growth, the soil moisture levels increase steadily. At approximately day 80, the crop coefficient peaks at a value of 1.2, which is subsequently met with a decline in the soil water content (Fig. 3.5, bottom). During this stage of peak water requirements from day 80 to 140, the crop enters its reproductive stage in which flowering and grain-filling occur. Concurrently, the water availability decreases as λ values decline near the end of the long rains and the relative soil moisture, s decreases from ca. 0.7 to 0.6. At the end of the growing season, the soil moisture levels slowly

Table 3.4: Summary of rainfall statistics with gauge record lengths greater than 40 years (n = 39) for short rains (SR, March through May), long rains (LR, October through December) or both seasons. We computed a modified Mann-Kendall test to designate significant trends ($p < 0.05$). Data source: CETRAD.

Parameter description	Season	Percent of stations with significant trends ($p < 0.05$)	Mean slope of all stations (Sen's method)	Standard error of slopes
Total rainfall [mm]	SR	8.75	0.2863	0.4469
Total rainfall [mm]	LR	1.25	-1.115	0.5240
Annual rainfall [mm]	Both	12.82	-0.0673	0.6194
Average depth per event α [mm]	SR	20.00	0.0654	0.0153
Average depth per event α [mm]	LR	21.25	0.0641	0.0180
Average event frequency λ [day^{-1}]	SR	33.75	-0.0010	0.0006
Average event frequency λ [day^{-1}]	LR	32.50	-0.0030	0.0005

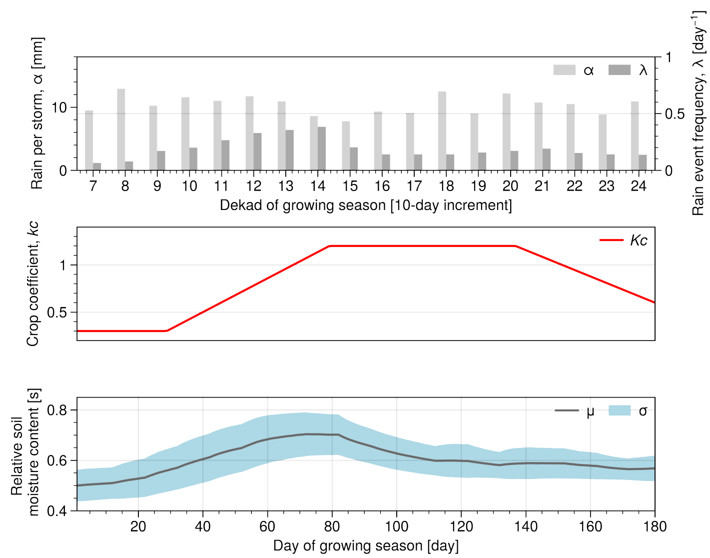

Figure 3.5: Three time-varying model parameters: 10-day average rain depth per event (α) and rain event frequency (λ) values starting with the planting date (March 2-11 in non-leap year) (top); daily crop coefficient (middle); and daily saturation (bottom).

decline and flatten until day 180 when the crop coefficient linearly decrease to a final value of 0.6. Interestingly, the crop coefficient and s align reasonably well over the course of the season. This alignment indicates that the Kc function that we use, which is derived from regional studies conducted by the FAO and previously described in section 3.2.2, works reasonably well for our study location as shown for a late maturing 180-day variety.

We then investigate the impact of stochastic rainfall on soil water availability for a single season. Figure 3.6a shows an example time series of simulated daily rainfall for a late maturing 180-day variety planted on day 60 using the Ol Jogi Farm rainfall climatology. As we expect, soil saturation increases in the early to middle part of the season (around days 50 to 100 in Fig. 3.6b), which aligns with the peak of the rainfall season Fig. 3.6a. The crop is moderately stressed over the entire season because the majority of the soil saturation time series falls between the stress and wilting points. In this particular simulation, the crop experiences the lowest levels of stress between days 60 to 80 whereas the highest levels of stress occur for the first 40 days of the growing

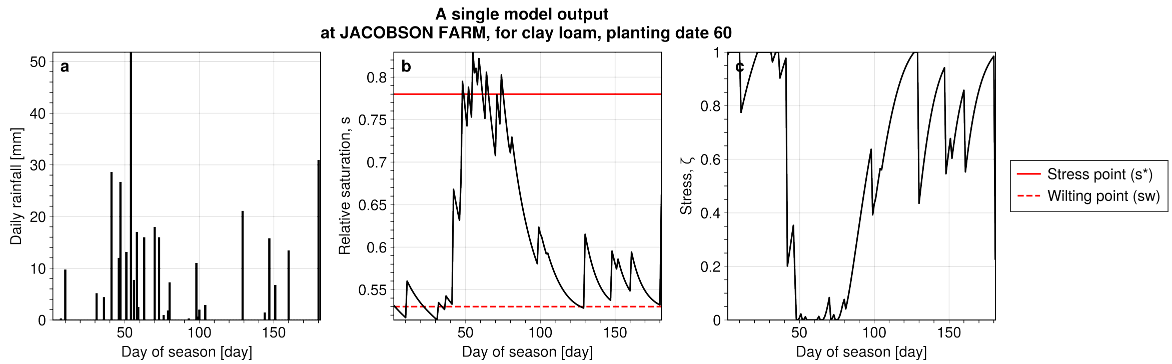

Figure 3.6: Time series of precipitation (a), relative saturation (b), and static stress (c) for a single simulation. Rainfall climatology is determined by Ol Jogi Farm station data, soil texture is clay loam, maize variety is 180 days, and planting date is day of year 60. Soil wilting point is red dashed line. Stress point (the point at which transpiration starts being reduced) is the solid red line. For a clay loam soil the stress point in units of saturation is 0.78 and the wilting point is 0.53.

season and between days 130 and the end of the growing season.

For the simulated 10,000 growing seasons, we investigate the distributions and time series of average soil saturation and static stress in Fig. 3.7. We find a seasonality in water availability in which soil moisture peaks between days (day of year) 135 and 145 for all simulations. The average soil moisture begins to decrease around day 145 before stabilizing around day 170. The simulations are prone to water deficit during the earliest and latest parts of the season in which the average saturation starts at 0.5 and ends the season around 0.55, on average. The crop is at the highest points of average static stress during these periods as shown in Fig. 3.7d.

The crop is generally stressed over the course of the season because the stress point is 0.78 and the saturation time series in Fig. 3.6b falls below and around the stress point for the majority of the season. The stress values of the crop ranges between 0 and 1, and in the time series, we see a larger increase in stress compared to the relative saturation time series. Additionally, we see greater variability in stress values compared to saturation values as the confidence intervals for stress are very wide especially in the later part of

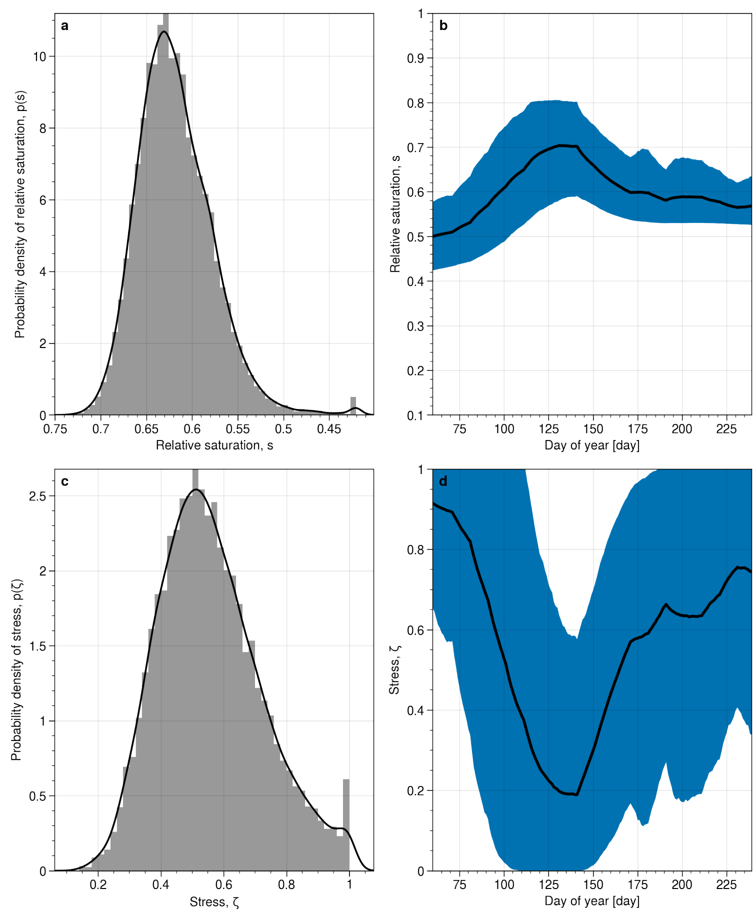

Figure 3.7: Probability distribution functions and time series of average soil saturation and stress with 90 and 10% confidence intervals for 10,000 simulations. Model parameters used are the same as those in Figure 3.6.

the season. This is due to the nonlinearity introduced in the conversion from relative saturation to stress.

3.3.2 Dynamic water stress

To convert average static stress into a yield metric that allows for the probability of crop failure, we use the dynamic stress equation. Dynamic water stress considers both the frequency and duration of excursions below the stress point and is used to estimate end-of-season yield as a fraction of the maximum possible yield for that variety as shown in equation 3.17. Calculating yield in this way is typical in the literature on numerical simulations of climate variability and crop yields (Van Ittersum et al., 2013; Roche et al., 2020). We show the PDF of dynamic water stress in figure A.2.

As shown in Fig. 3.8 for the 180-day variety, the relationship between rainfall and yield is not a linear one, but rather is asymptotic. We also include the relationships between seasonal rainfall and yield for the early, medium, and late maturing varieties categories in A.4. As the crop experiences more rainfall, yields increase up to a ceiling which is defined by the yield potential of the cultivar. We find that yields are closer to the maximum potential yield with seasonal rainfall totals greater than ca. 600 mm. When the seasonal rainfall totals are less than approximately 450 mm, the possibility of crop failure is introduced. We find that seasonal rainfall totals between ca. 200 and 450 mm, there is a large range in possible outcomes for farmers. For a given seasonal total within this range, we find that farmers can reach up to 60-75 percent of the maximum possible yields while others experience total crop failure. This relationship also holds true for early and medium maturing varieties as shown in A.4. For seasonal rainfall totals below 200 mm or so, the majority of simulations result in total crop failure. Therefore while the seasonal total of rainfall is important, there are also characteristics of within season

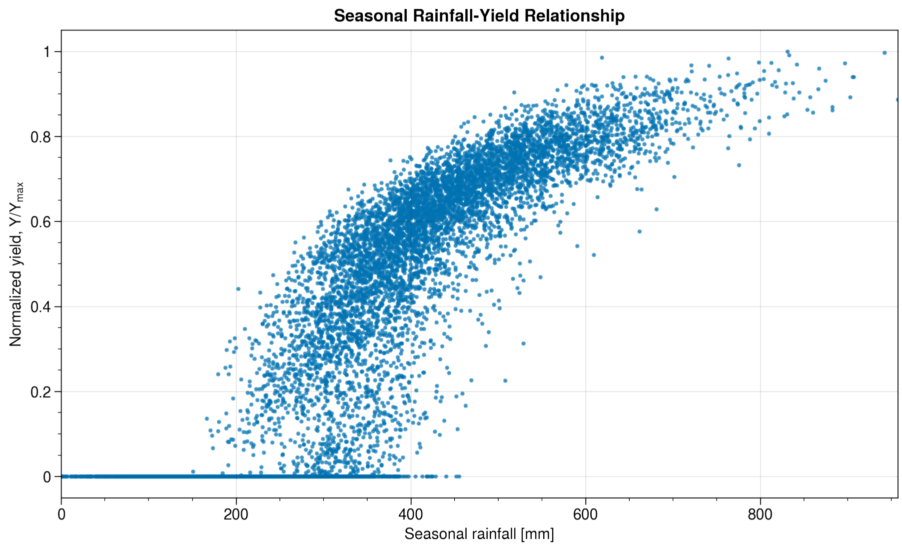

Figure 3.8: Scatterplot of seasonal rainfall and yield for 180-day maize and study site conditions. End of season yields are normalized by the maximum yield for all 10,000 simulations: 2.8 t/ha.

rainfall that have a pronounced effect on yield outcomes.

3.3.3 Cultivar choice mediates probability of crop failure

We then investigate the effect of cultivar choice (i.e. LGP) on yield and crop failure using the methods described in section 3.2.3. Figure 3.9 shows the joint probability distribution of yield and rainfall for three categories of maize varieties: early, medium and late maturing. The average rainfall and yield for each category of maize varieties is denoted as a black "x". In Table 3.5, we provide the average statistics that correspond with Fig. 3.9. We find that the average seasonal rainfall increases in proportion to the length of the growing season for each of the maize varieties categories. Because early maturing varieties require less time to grow (e.g. 80-110 days) compared to late maturing varieties (e.g. 150-180 days), they receive less average seasonal rainfall.

Despite having lower seasonal totals due to their shorter maturity period, early-maturing crops are more likely reach their maximum potential yields while medium and

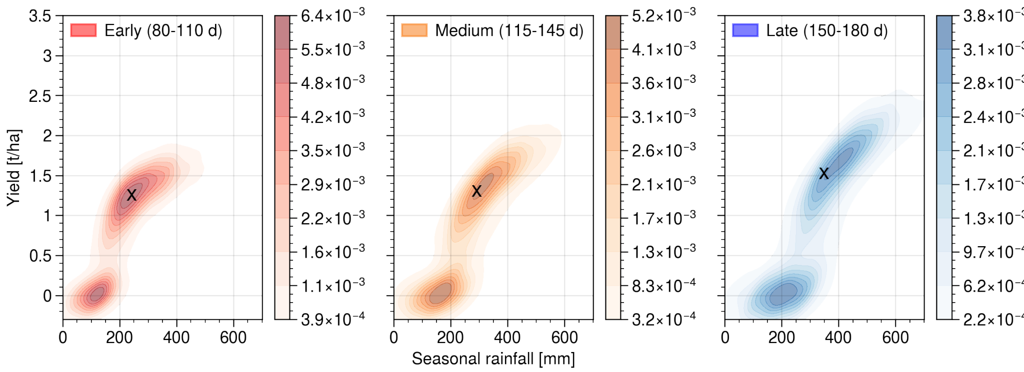

Figure 3.9: Joint probability distribution of three maize varieties categories using Ol Jogi rainfall climatology. The average rainfall and yield for each subplot is denoted with 'x'. Non-zero values of yield were used to calculate this average.

Table 3.5: Summary statistics for three maize variety categories. These results are from the methodology described in section 3.2.3 and explained in section 3.3.3. Average seasonal rainfall is increasing in proportion to the length of the growing period (LGP) in which early maturing varieties take between 80-110 days to grow; medium maturing take 115-145 days; and late maturing take 150-180 days.

Statistic	Early	Medium	Late
Average Seasonal Rainfall (mm)	221.6	270.8	328.1
Average Yield (t/ha)	1.18	1.23	1.45
Crop Failure (%)	26.77	31.97	34.94

later maturing crops only gain a fraction of their maximum yields (A.4). Additionally, under low seasonal rainfall totals, medium- and-late maturing crops have a higher incidence of crop failure, which is indicated by the spread of probability densities around 0 yield between 0 and 300 mm of rainfall in Fig. 3.9. In comparison, early-maturing crops are less likely to fail, which is indicated by the less dense probabilities below 200 mm of rainfall. These trends are reflected in the percentage of simulations that resulted in crop failure: approximately 27%, 32%, and 35% for early-maturing, medium-maturing, and late-maturing, respectively, as shown in Table 3.5. Late-maturing varieties also have a larger spread in potential yields compared to early- and medium-maturing varieties.

3.3.4 Regional rainfall climatology trends

We investigated interannual trends in seasonal totals, storm depth and inter-storm arrival rate for the two rainy seasons. Overall, we find that changes in seasonal totals are minimal across the stations, shown in Table 3.4, and as visualized for Jacobson Farm in Figure 3.2. However, we do find significant trends for increasing intensity, α, and decreasing frequency, λ in both rainfall seasons ($p < 0.05$, Table 3.4) using the modified Mann-Kendall statistical test. Comparatively, total rainfall for both seasons and annual rainfall shows a muted change with fewer stations showing significant trends. Our results are consistent with those of Franz et al. (2010), which analyzed a less recent dataset of 11 stations in the same region. In Figure 3.2, we use the Jacobson Farm rain gauge which has the longest record (79 years) to show the shifts in rainfall processes that occur in this region. Here, we see that while total rainfall for either season does not change significantly, we see an increase in α and decrease in λ over the period ($p < 0.05$, Fig. 3.2). This indicates that storms are becoming more intense and less frequent. The relevance of this interannual variability in rainfall is further discussed in section 4.5.

3.3.5 Long-term trends in yield and crop failure using the Jacobson Farm rainfall climatology

After changing 10-day average rain depth per event (α) and rain event frequency (λ) values to represent those in an earlier era (1930s) and present day (2010s) at Jacobson Farm, we show how cropping outcomes have changed over the 80 year record in Fig. 3.10 and Table 3.6. During this period, we see the rain depth per event (α) and rain event frequency (λ) change where the average α values increased from 5.71 mm per storm event in the 1930s to 10.99 mm per storm event in the 2010s. Conversely, the average λ values decreased from 0.35 rainfall events per day in the 1930s to 0.16 rainfall events per day

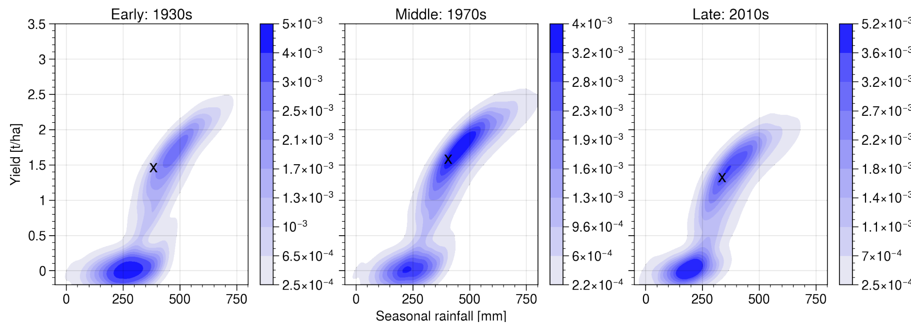

Figure 3.10: Historical change in seasonal rainfall and yields for Jacobson Farm climatology. We defined three eras of rainfall conditions in which we set the rain depth per event (α) and rain event frequency (λ) parameters based on the relative trend line in Figure 3.2. The average rainfall and yield for each subplot is denoted with 'x'. Non-zero values of yield were used to calculate this average.

in the 2010s. We see a marked shift in the average seasonal rainfall generated from the simulations, which exhibits an inverted U-shape. There is an increase in average rainfall between the 1930s and the middle of the record from 362 mm (standard deviation, SD, 139 mm) to 385 mm (SD 154 mm), and a decrease to the later part of the record in the 2010s of 316 mm (SD 136 mm). The average seasonal rainfall total in the latest part of the record was dramatically lower than the earlier two data points.

Interestingly, average crop yields also follow the inverted-U shape: At the start of the record, average crop yields were the lowest at 0.72 t/ha. Crop yields increased to the maximum of these three eras in the middle of the record to 1.05 t/ha and then decreased again to 0.84 t/ha in the 2010s. Crop failure rates show a similar pattern in which the crop failure rate was lowest in the middle part of the record (31%) and highest during the early part of the record (48%) and the present day crop failure rate is in between at 33%. The coefficient of variation of rainfall increased from 0.38 in the 1930s to 0.43 in the 2010s.

3.4 Discussion

Table 3.6: Summary statistics for historical trends in crop failure and yield. Parameters set to Jacobson Farm with a planting date of day 60 and a 180 day maize variety.[a]

Metric	Early: 1930s	Avg: 1970s	Late: 2010s
Average Seasonal Rain Depth (α) (mm)	5.71	8.33	10.99
Average Seasonal Rain Frequency (λ) (day^{-1})	0.35	0.26	0.16
Average Seasonal Rainfall (mm)	362.3	384.6	316.1
SD Rainfall	139.2	154.3	136.0
CV Rainfall	0.38	0.40	0.43
Average Yield (t/ha)	0.72	1.05	0.84
Probability of Crop Failure (%)	48.26	30.92	32.80

[a]SD, standard deviation. CV, coefficient of variation.

3.4.1 Rainfall trends and impacts on maize production

Warming global temperatures directly cause interannual variability in rainfall—i.e., the shift towards fewer and more intense storms (IPCC, 2007)—as observed at Jacobson Farm and other gauges in the field site (Table 3.4). With air temperatures and atmospheric water vapor rising in East Africa as well as globally, extreme precipitation events are expected to increase (Trenberth, 2011; Solomon et al., 2007). We find that while rainfall seasonal totals are not changing, the average rainfall event depth and frequency have changed in several sites in Laikipia, Meru and Nyeri counties.

When considering the extremes of the 80-year rainfall record (i.e. 1930s versus 2010s rainfall climatology), we see changes in cropping outcomes among maize farmers. Overall, we find an inverted-U shaped relationship from the three time periods of interest where the average yields peaked and the probability of crop failure was lowest in the middle of the time series. We found that the average seasonal rainfall shifted from a period of high seasonal rainfall in the 1970s and lower seasonal rainfall in the 2010s. Our results indicate that the environmental conditions for growing maize were most difficult in the earlier part of the record, they became easier in the 1970s, and now are becoming more difficult again. The reasons for these changes stem from the interactions between shifts in mean seasonal rainfall, increasing storm depths, and decreasing storm even frequency. It is notable that despite the very low seasonal rainfall totals in the 2010s, the average yield is not lowest and the crop failure rate is not the highest seen across the three eras. The lack of correlation between average seasonal rainfall and crop yields highlights the nonlinearities between seasonal totals, rainfall statistics (average depth and frequency) and yield that interact and make it difficult to predict yield using a single metric. Previous work has demonstrated similar findings such that the intensity and duration of individual rainfall events is important for hydrological partitioning of precipitation in addition to annual or seasonal totals (Taylor et al., 2013; Apurv et al., 2017; Singer and Michaelides, 2017; Kipkemoi et al., 2021). We also find an increase in the coefficient of variation in rainfall showing that farmers experience increasingly variable rainfall conditions.

Previous studies have sparked concern over the future of rainfall in East Africa and formed a consensus about the role of climate change in influencing rainfall patterns (Nicholson, 2017; Shongwe et al., 2011). Specifically, these studies have demonstrated negative trends in the magnitude of the long rains (March-May) in East Africa and the Horn of Africa Drylands (HAD) (Lyon and DeWitt, 2012; Liebmann et al., 2014; Williams and Funk, 2011; Funk et al., 2018b, 2019). Other studies have shown regional trends towards more extreme rainfall events (Harrison et al., 2019), particularly for the short rains (October-December) (Shongwe et al., 2011) and in central Kenya (Schmocker et al., 2016). In addition to decreased seasonal totals, trends towards more extreme rainfall events will likely contribute to decreased end of season yields and increased rates of crop failure due to the timing and nature of rainfall and crop water requirements of maize over the season. Increasingly heavy rainfall events pose a threat to crop production due to extreme surface runoff and subsequent erosion or from flood events (Liniger and Thomas, 1998; Liniger and Weingartner, 1998).

Small-scale producers depend on reliable rainfall. We show that maize yields are lower in the present day in comparison to yields 40 years ago due to reduced seasonal rainfall totals and increased rainfall variability. In order to attain end of season yields that are profitable, maize growers must avoid both sodden conditions due to extreme rainfall and drought conditions caused by dry spells (Rigden et al., 2020). Our results show that despite having seasonal rainfall totals that should be adequate for maize growth, the end of season yields and chance of crop failure are highly variable. This is because the crop coefficient of maize does not necessarily align with the periods of highest rainfall when planting is conducted in early March and in these systems where irrigation may be absent. For this reason, hybrid maize has been developed to withstand variable rainfall and drought during the growing season. We also find that intraseasonal rainfall variability has a pronounced impact on farmer outcomes for varieties with different maturity periods.

3.4.2 Cultivar choice moderates exposure to stress

Varieties with shorter maturation periods reduce their exposure to water stress and fail less often compared to varieties with longer maturation periods. We show that modest decreases in relative soil saturation lead to dramatic increases in the water stress of simulated maize. When the length of the growing period is shortened for an early maturing crop to reach the flowering and grain-filling phenological stages faster, there is a reduced chance of exposure to water deficits. Therefore, we find that early maturing varieties have a shorter exposure to rainfall variability and therefore are less likely to fail whereas long-maturing varieties grow for longer periods and are more susceptible to longer periods with no rainfall. These early maturing varieties escape longer dry periods and therefore may be a better option for small-scale maize producers under increased levels of rainfall variability.

Cultivar choice is one of the primary adaptation strategies that farmers can control to minimize the stress experienced by the crop. We show that early maturing crops are the best choice given their fast maturity period and lowest probability of crop failure. Late maturing crops will produce higher yields in the simulations that do not fail, however, they are exposed to longer periods of no rainfall due to taking more time to grow and therefore have a higher risk of crop failure. Early maturing varieties are the least likely to fail just on the basis of requiring less time to grow and thus have a shorter exposure to the rainfall process. Furthermore, early maturing varieties are often bred or have been genetically modified to be drought tolerant. Drought tolerance is often developed for a specific set of environmental conditions and thus can be hyper-localized.

In a survey of 500 East African farming households, Erenstein et al. (2011) found that the most desired attribute of maize varieties was yield potential followed by early maturity. These characteristics were perceived to be more important than drought tolerance. Maize cultivars can exhibit one of two traits: drought avoidance as in the case of early maturing varieties and drought tolerance in the case of hybrids that are bred to tolerate reduced available water. Early maturing varieties are considered drought avoidant because they are thought to complete their most drought-sensitive stage (flowering) before a drought occurs such as at the ending of a growing season (Barron et al., 2003; Morris, 2001).

We show that by growing a late maturing variety, the crop is still subject to terminal drought and crop failure due to the timing of rainfall throughout the season. However, a trade-off exists in selecting a late maturing variety. As shown in the empirical data of maize cultivars and yield (Figure 3.3), there is an opportunity for higher yields for late maturing maize varieties. In cases where farmers do not have access to irrigation and would prefer a crop that has a greater chance of success despite a potential yield penalty, farmers may prefer early maturing crops (Barron et al., 2003).

There are two possibilities that plant breeders and farmers can elect for: to optimize for survival by minimizing the variance of stress experienced by the crop or to optimize for greater biomass in order to get the maximum yield. In light of this, farmers may be interested in maximizing their yield as well as limiting crop failure by planting a medium maturing variety. However, in areas with high rainfall variability there might be reason to plant only early maturing varieties either with the intent of minimizing crop failure or in order to double crop within the season. In these settings where a farmer's goal is to prevent crop failure, early and extra early maturing varieties are more effective for a planting date of March 1 as described in our model. There may be a role for varieties with other maturity durations when using different planting dates.

3.4.3 Declining yields, increasing crop failure rates and household-level impacts

The increased variability in rainfall and decreased seasonal totals have serious implications for agriculture in East Africa and other regions of the African continent where large portions of the population suffer from food insecurity (Funk and Brown, 2009). These changes in mean annual rainfall and increased storm size may make it increasingly difficult for farmers to practice agriculture as they have done so in the past. As shown in the temporal analysis of rainfall trends, the timing and distribution of rainfall has changed within one to two generations of farmers.

Climate change and climate variability impacts can shock the economic system by altering food prices and so affecting food demand, nutrition, and human livelihoods (Herrero et al., 2010). We have already shown that season-to-season variability in rainfall is high, and the distribution of rainfall in addition to the total is of paramount importance to Kenyan agriculture. In Kenya, declines in per capita maize production have been reported for certain regions (Funk et al., 2018a). Much of this change is due to variable seasonal rainfall and the incidence of crop failure. Our results are consistent with work that projects reduced rainfed maize yields in Kenya (Herrero et al., 2010; Thornton et al., 2010). While yield gains have been projected for certain highland areas in the temperate areas of Kenya (Thornton et al., 2010), farmers in the semiarid and arid lowlands are predicted to experience diminished yields, likely forcing them to sow varieties with shorter maturity periods.

In order to capitalize on any potential yield increases in the highlands and reduce as much as possible the decreased yields in the majority of Kenya, further investment in new varieties, improved inputs, and services will be needed (Herrero et al., 2010; Hansen et al., 2011). Investing in varieties that conserve root-zone soil moisture or have deeper roots may be especially important in a warming climate (Rigden et al., 2020). Access to irrigation and water harvesting more generally will be an important way for farmers to buffer negative climate impacts. In this region in central Kenya, access to irrigation resources is not ubiquitous and even those farmers with access to irrigation experience high spatial and temporal variability in its availability (Gower et al., 2016). McCord et al. (2018) show that farmers in the Mount Kenya region with greater relative variability in water flow from irrigation are more likely to uptake adaptation measures such as choosing new seed varieties. This is a positive indication that perhaps the agriculturalists who are more impacted by rainfall variability due to reduced irrigation access are likely to employ adaptation measures or at least experiment with those measures such as changing to an early or extra-early maturing varieties (McCord et al., 2018).

3.4.4 Study limitations and future research

We made some important assumptions in our model, a common practice in such studies (e.g. Challinor et al. (2009); Tesfaye et al. (2016)). First, other than the rainfall climatology and the crop coefficient all other variables were constant. We do not simulate the effect of inputs such as fertilizer application or irrigation use. The model assumes that nutrients like nitrogen are available in adequate quantities that do not limit growth crop and yield. Second, we do not simulate other crops which might be intercropped with maize (e.g. beans, potatoes), and we do not simulate crop rotation or varying cropping densities. Additional studies may begin with any of these assumptions to further evaluate their effects on maize production.

Future studies would benefit from adding empirical agronomic or decision-making data as inputs or validation datasets. Our study has shown that available water content is lowest during the latter part of the season when the crop coefficient is the highest and the rainfall slackens. A follow-up study could investigate the intra-seasonal nature of water stress to demonstrate what duration of stress during the season makes the largest impact in terms of yield. To answer this question, an empirical dataset is needed that includes both intra-seasonal stress dynamics and end of season yields. Additionally, we did not constrain the behavior of early, medium, and late maturing varieties other than changing the length of their growing periods. Early maturing maize is often bred or genetically modified to be drought tolerant and thus should reduce the probability of crop failure. These constraints could be added as parameters in the dynamic water stress calculation or by altering the crop coefficient for hybrid maturities. Field-collected data or on-farm trials would be needed to constrain these parameters.

3.4.5 Conclusions

This stochastic ecohydrological model represents conditions that will become more common as climate variability and climate change alters rainfall in tropical and semiarid systems. By using historical rainfall to generate stochastic conditions for average depth and probability of rainfall we simulated a dryland environment for small-scale producers. We considered a common crop choice (maize), soil type, and planting decisions (i.e. timing of planting) that represents hundreds of millions of small-scale producers in regions vulnerable to climate change and variability. We investigated the role of rainfall variability in explaining current and past agricultural outcomes such as yield and likelihood of total crop failure. Additionally, we show the importance of cultivar choice in determining yield potentials and the vulnerability of late maturing varieties to rainfall variability. The large divergence in farmer outcomes at low seasonal rainfall totals is concerning. The within season characteristics of how and when rainfall occurs creates the water stress environment that leads to crop failure in some cases and relatively high yields in others.

How we characterized rainfall variability and its relationship to crop phenology (crop coefficient) was a novel contribution to a field where hydrologic processes are often considered separate from the phenology of the crop. We defined water availability as a function of stochastic rainfall and soil parameters whereas the crop coefficient governed water demand. Our model especially considers the rainfall depth and frequency (α and λ) parameters as forces that drive stochastic rainfall. Thus we can simulate time-varying soil moisture which changes over the course of the season due to shifts in the water required by the crop and changing rainfall statistics.

In the face of climate change, where direct changes in water availability, temperature, and increased prevalence of pest and diseases are possible, farmers can adapt through two key management strategies: choice of cultivar maturities and planting dates (Van Ittersum et al., 2013). Farmers need to select locally relevant planting dates and cultivars with appropriate maturities to minimize crop failure. Future study on the role of planting dates would illuminate the relative advantages of early to late maturing crops planted during Kenya's two primary rainy seasons. With appropriate climate data and locally-relevant agronomic conditions, this type of modeling can be used as a heuristic to improve our understanding of the impacts of climate variability on farming outcomes in other contexts.

Chapter 4

Modeling the spatial location of cellular towers in Zambia as a Gibbs point process

Abstract Mobile broadband coverage is not equitably spread across the globe. Improving telecommunications access for those who are underserved can lead to many social benefits such as improved access to health information, economic opportunities, and emergency alerts. Coverage in remote locations in the Global South is not fully understood due to lack of data on cell tower locations. Our study aims to better understand the spatial distribution of telecommunications access in Zambia: a context where the majority of the population lacks access to electricity and there is a growing divide between urban and rural residents. We model the location of cell phone towers, a proxy for broadband coverage, in such a way that allows for the inclusion of environmental covariates and specific forms of interaction between cell tower points. To do this, we use a Gibbs model which decomposes the conditional intensity of the spatial point process into trend ("first-order") and interaction ("second-order") components. The first-order com-

ponent is made of four covariates (distance to roads, livelihood zones, power plants, and population density) and the second-order component are various inter-point interaction functions that describe inhibition and clustering processes. We found that while towers tend to be clustered in urban areas, distance to highways is the most important covariate that influences tower placement in rural areas. Our work highlights the disparities in telecommunications access for rural Zambians who also lack access to electricity. Future studies can apply our methodology to study cell tower placement and uncover spatial dynamics of telecommunications access, which has implications for applied fields such as early warning systems, public health, and international development.

4.1 Introduction

Improving access to mobile phones and mobile internet is a key way to make fast progress on many sustainable development goals (Rotondi et al., 2020). Specifically, well-located cell phone towers and strong mobile broadband coverage are imperative for the diffusion of mobile phone and mobile internet services (Baumüller, 2018; Krell et al., 2020). However, a widening gap in telecommunications access exists between rural and urban residents in the Global South. Those unconnected to mobile phone and internet coverage tend to be poorer, less educated, and remotely located (Bahia and Suardi, 2019). Nearly 600 million people on the African continent lack access to electricity (World Bank, 2018b), which is both necessary to charge mobile phones and for a grid connection to power cell tower infrastructure. In order to quantify disparities in telecommunications access, spatial point process models can be used to study patterns in point-based infrastructure such as cell towers. Several socio-environmental factors may influence the location of cell towers such as road networks, differences in wealth, livelihood, and electrification, and population density that can explain disparities in access. Due to systematic and prevailing inequalities, individuals located in rural areas and who are disconnected from road networks are often left without access to information and communications technologies (ICTs) (Nakasone et al., 2014; Baumüller, 2018).

Mapping the location of cell towers and correlating areas that are unserved with population data is very useful (Song, 2017), however, this is difficult because publicly available datasets of verified cell tower locations are practically non-existent. While mobile phone coverage maps can be available, external validation of cell tower locations are rare and this information is unlikely to be available in the public domain. This is because telecommunications companies treat information related to infrastructure, specifically cell tower locations, as sensitive because it improves their market position (Fida and Marina, 2019). If companies do release information they make it subject to a non-disclosure agreement and may only make it available to policymakers and regulators (Fida and Marina, 2019). Thus, researchers and development practitioners must turn to crowdsourced measurements of cell tower locations.

OpenCellID is a large-scale quasi crowd-sourced dataset of cell towers globally. A growing body of literature is centered around improving estimates of cell tower locations obtained by crowdsourced data in which the sample collection process is uncontrolled. Algorithms have been developed to improve the accuracy of crowdsourced tower locations, especially with the goal of seeing what the reach of mobile infrastructure is in developing countries. Fida and Marina (2019) provide a review of the pitfalls of existing algorithms that are used to localize cell towers from crowdsourced estimates, and overall conclude that all of the methods are inadequate when verified cell tower locations are not available. Currently, the only way to access cell tower information is through crowdsourced data from user devices and then make inferences based on estimates of the actual locations (Fida and Marina, 2019). Using cell tower data that was contributed by users is helpful because it is cost-effective; however, attaining accuracy and robustness is challenging because the measurements are uncontrolled, i.e. multiple devices report at various times and from random locations. To make improved estimates of cell tower locations, a ground-truthed dataset would be needed which is rare.

Spatial analyses that provide estimates of the influence of environmental covariates on the spatial pattern of cell towers in the Global South are useful despite the lack of fully verified cell tower locations. Stochastic point process models such as the Markov ("Gibbs") point process model where the points (e.g. location of cell towers) are discrete random variables are instructive for spatial analyses (Baddeley et al., 2015). Gibbs models essentially provide a regression framework that uses point data as the input and yields first order effects which estimates the influence of covariates on the trend, and second order effects: a representation of the interaction between points (Sweeney and Gómez-Antonio, 2016). The inter-point interaction may be an inhibition process which prevents two events (i.e. cell towers) from being located within a certain distance to each other. Alternatively, the inter-point interaction could exhibit clustering in which points are attracted to each other. A Gibbs model is appropriate for studies of the diffusion of infrastructure and industries because because they solve problems with identification, spillovers, and provide estimates of which covariates most effectively describe the point pattern (Sweeney and Gómez-Antonio, 2016).

The overarching goal of this research is to understand the influence of the trend and inter-point interaction on the spatial point pattern of cell tower locations in Zambia. Understanding this is critical to making access to mobile phones and mobile internet universal, which yields social benefits such as the diffusion of government services and capital and wealth distribution. We focus our analysis on one cell tower company in Zambia: Zamtel. Specifically, we ask the following research questions for our study area:

1. How do socio-environmental variables relate to the location of cell towers? Specif-

ically, is each covariate (population, roads, electricity, and livelihood zones) more or less likely to be found near a cell tower?

2. After accounting for variation in the cell tower intensity due to the covariates (i.e. first order effects), what processes govern the inter-point interaction between cell tower locations? I.e. is there evidence of clustering or inhibition between points?

To answer these questions, we investigate whether the intensity of the point pattern is dependent on the environmental covariates. Then, we allow for the effects of the environmental covariates on the intensity before studying the interaction between the points. The workflow for fitting the models are as follows. First, we formulate a Gibbs model with four environmental covariates. Next, we fit the models to the point pattern dataset and specify irregular parameters. Then we interpret the findings and lastly we identify the weaknesses of the model fits using validation tests and diagnostics.

Our paper is structured as follows: Section 4.2 provides an overview to Zambia as a case study. Section 4.3 describes our point pattern dataset: both the cell tower points and the environmental covariates. Section 4.3 also contains our methods namely the Gibbs models and the approaches used for model specification testing and diagnostics. We then apply the methods to cell tower locations in Zambia. In section 4.4 we present our results in terms of the point pattern description, the Gibbs model estimates, and validation and diagnostics. Lastly, we present our discussion of our main findings, the study's limitations, and their implications in section 4.5, and we summarize our main conclusions in section 4.6.

4.2 Case study site: Zambia

We focus our study on Zambia, which is a landlocked country in southern Africa (Figure 4.1) that has been undergoing rapid development since the early 2000s (Bayliss and Pollen, 2021). Zambia is sparsely populated with 55% of its population residing in rural areas and the rest residing in urban areas (World Bank, 2018b). Zambia is composed of savanna drylands which face a strong seasonality in rainfall in addition to intra- and interseasonal rainfall variability (Vergopolan et al., 2021). The majority of the land is occupied by small-scale farms that are smaller than 5 ha (Jayne et al., 2016). The make-up of agricultural production is 35.8% small-scale farms, 53% medium-sized farms (5-100 ha), and the remaining is large scale farms (Jayne et al., 2016). Zambia consists of over 100 districts and can be grouped into 21 livelihood zones (FEWS NET, 2014). The industry that contributes the most to Zambia's Gross Domestic Product (GDP) is copper mining; however, the primary livelihood of Zambians is farming. Hydroelectric power generation dominates Zambia's energy generation with 85% (2,380 MW) of the 2,800 MW of installed electricity generation capacity coming from hydroelectric dams (USAID, 2020). The remaining make-up of Zambia's generation capacity comes from coal, heavy fuel oil, and a small portion of solar power (USAID, 2020). Only 4% of Zambia's rural population has access to electricity compared to 67% of the urban population having access to electricity. An estimated 7.2 million households lack electricity (USAID, 2020) out of approximately 18 million people in Zambia (World Bank, 2018b).

4.3 Methods and data

A point pattern is an enumeration of events (i.e. points or objects of interest) in a defined study area. Points can be located in either two- or three-dimensional space. For

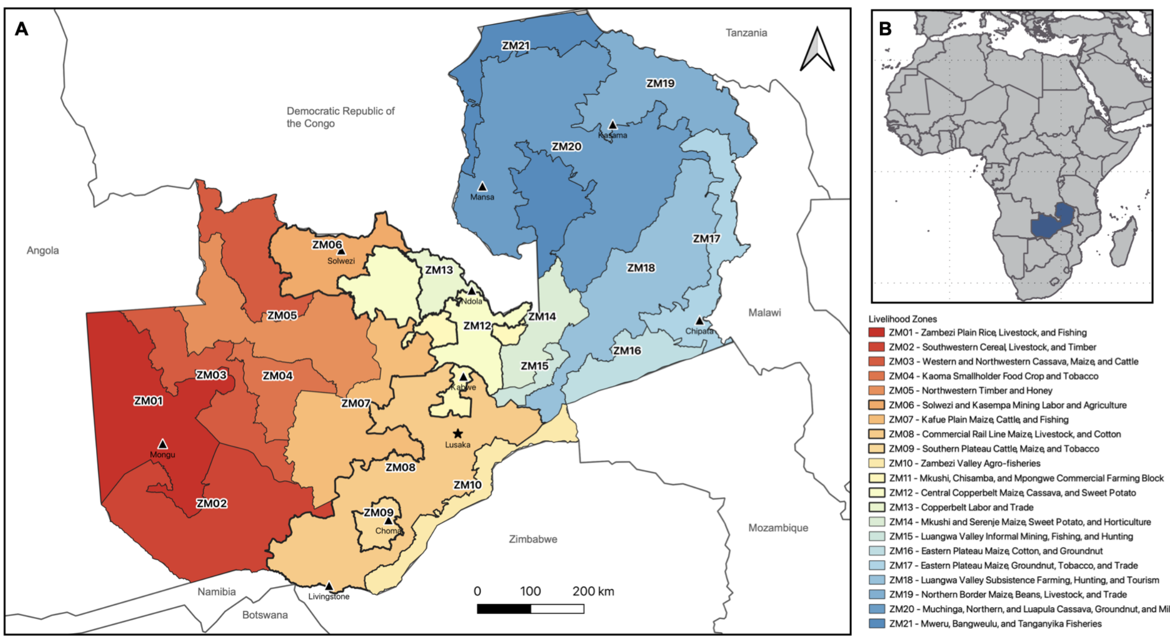

Figure 4.1: Map showing study site: A) Zambia's 21 livelihood zones as presented in (FEWS NET, 2014). Livelihood zones 6, 8, 9, 11, 12, and 13 are outlined and they represent the mining, commercial farming, and more urban areas of Zambia. B) Location of Zambia on the African continent.

Table 4.1: Description of environmental covariates

Variable	Source	Year Published	Spatstat Object Type
Roads	RCMRD[a]	2015	Image (Distmap)
Livelihood Zones	FEWSNET[b]	2014	Image (Categorical)
Population	Worldpop	2017	Image
Power Plants	KTH[c]	2015	Image (Binary)

[a]Regional Centre for Mapping of Resources for Development.
[b]Famine Early Warning Systems Network.
[c]Royal Institute of Technology Stockholm.

this study, we consider cell tower locations as the observed point pattern x which is a realization of a random point process $\mathbf{X}$ in a planar 2D space. We adapt the mathematical notation used by Baddeley et al. (2015) and Sweeney and Arabadjis (2021). A point pattern is a collection of point coordinate locations, $u = (u_1, u_2)$, in $\mathbb{R}^2$. The number of points is $n(x)$ and the vector of points is noted by $\mathrm{x} = \{x_1, x_2, ...x_n\}$. We observe the point pattern x in the window W. Additionally, we consider subregions, $B \subset W$ in which a count of points exists in a subregion, $n(\mathrm{x} \cap B)$. The intensity, λ, is the density of points per area, which is defined $\lambda = n(\mathrm{x})/|W|$.

Point processes have first- and second-order properties. First-order properties relate to the density (i.e. intensity) of events across the study area. Second-order properties are related to the spatial dependence (i.e. interpoint interaction) of the points. We characterize the trend and interaction of the point process against Complete Spatial Randomness (CSR). CSR is a homogeneous Poisson process in which there is no spatial trend and there is no interaction between points. Additionally, the number of points that fall in a subregion, B, will be Poisson distributed with the mean: $\mu = \lambda|B| = \mathbb{E}n(\mathrm{x} \cap B)$.

4.3.1 Trend and interaction

The variation in the density of points within W is the **spatial trend**. This is the first-order variation or mean effect. We use kernel smoothing of the point density in which a density estimate $\hat{\lambda}(u)$ is the weighted sum of the nearby per-pixel densities. The second-order effects are the **interpoint interactions**. The interaction relates to the covariance of the point process such that a positive covariance will show a pattern in which points are attracted to one another, i.e clustered. In the case of clustering, the number of points next to any one point is relatively high. If points exhibit inhibition then there is negative covariance. In the case of inhibition, the count of points nearby one point will be relatively low. The relative high or low count of nearby points is in relationship to CSR: a baseline process of no interaction (Sweeney and Arabadjis, 2021).

To begin, we define the K-function of point process X as the number of expected neighbors of a point divided by the intensity:

$$K(r) = \frac{1}{\lambda}\mathbb{E}[t(u, r, \mathrm{X} \mid u \in \mathrm{X}]$$

where the count of points in X within r distance of a point u is $t(u, r, \mathrm{X})$. We use a transformed version of the K-function to measure interaction: $L(r) = \sqrt{\frac{K(r)}{\pi}}$ or $L(r) - r$. This transformation results in plots that are easier to interpret; deviations from CSR (=0) indicate clustering (>0) or inhibition (<0) (Besag, 1977).

4.3.2 Spatial process models

We define stochastic models for the point process that include environmental covariates, to capture first-order effects, and an interaction term to capture second-order effects. The second-order effects are specified using specific functional forms for the point processes' covariance that result in a coherent model. While several frameworks for modeling point processes exist, one of the most widely adopted and well supported by software are Gibbs process models. Details of probability theory underlying Gibbs models is available in Moller and Waagepeterson (2003) and applied modelling and software are reviewed by

Baddeley et al. (2015). We provide a sketch of the framework below.

Gibbs models can be expressed in multiplicative in form, $\lambda(u, x) = \beta(u)\gamma(u, x)$ such that the conditional intensity of the process ($\lambda(u, x)$) at location u is the product of a trend component at u, $\beta(u)$, and interaction at location u as a function of nearby points in x, $\gamma(u, x)$. In this framework, parameters β and γ appear in the likelihood and are termed "regular", and can be estimated using methods such as maximum pseudo-likelihood or full maximum likelihood using Markov chain Monte Carlo (MCMC). The interaction component also will include what are termed irregular parameters (such as radius of interaction), and those are estimated outside the model using profile likelihood. As noted above, the outcome modeled is the conditional intensity of the process, $\lambda(u, x)$. This allows for feasible estimation and avoids complex normalizing constants that appear in the intensity. The conditional intensity is the intensity at location u given the rest of the point pattern $\mathbf{x}$: $\lambda(u|x)$.

We use the package **spatstat** in R, which provides a robust framework for Gibbs process models and is often used in scientific applications (Baddeley and Turner, 2005). The software also encodes a suite of methods to assess model fit and validation of the trend and interaction. Similar to any regression modeling context, the goal is model variation in the outcome (the point process in this case) as a function of predictors. The additional complication in this setting is the trend and interaction components are estimated simultaneously.

We originally evaluated four functional forms for the interaction component of the Gibbs point process model: Strauss Hard-Core (SHC), Soft Core (SC), Diggle-Gratton (DG), and piecewise constant (PC). Our expectation prior to working with the data was that the process, conditional on adequately modeling the trend, would exhibit repulsive interaction. This would be expected in a either a competitive environment or from the perspective of optimal allocation of scarce resources. After working with the data, it became clear that the process is clustered, conditional on trend. We then moved on to evaluate forms of interaction that allow for clustering: Geyer-Saturation (GS) and piecewise constant (PC). As a baseline, we evaluated against the imhomogeneous Poisson process which includes trend but no interaction, $\lambda(u,x) = \beta(u)$. We fit the models using the loglinear form $log\lambda(u|x) = \eta^T Z(u) + \psi^T T(u,x)$. We recover the estimated trend, $\hat{\beta}(u) = \exp(\hat{\eta}^T Z(u))$, and interaction, $\hat{\gamma}(u,x) = \exp(\hat{\psi}^T T(u,x))$, components. We include the sub-cases for the inhomogeneous Poisson process, $\lambda(u,x) = \beta(u)$, and homogeneous Poisson process, $\lambda(u,x) = \beta$, in this framework. When we interpret the model estimates we consider the trend effects with respect to the environmental covariate which effects the conditional intensity at location u. The interaction component is a measure of the pairwise proximity between points such that a point with a distance r of u is a multiplicative factor of $\hat{\gamma}(u,x)$. A value equal to one is no effect, greater than one is clustering, and less than one is inhibition. Each of the functional forms of interactions define different weights to points within r distance of u.

Environmental covariates

Environmental (spatial) covariates can be a spatial function $Z(u)$ that are defined at all spatial locations, u, of the point pattern in W or they can be another spatial point pattern. Covariates can be either continuous or discrete. We work specifically with four covariates—roads, population density, livelihood zones, and power plants—found in Table 4.1. The steps used to create the covariates follow:

Roads The roads are primary, secondary, and tertiary roads which provide connectivity between urban areas and rural towns. For roads, we need a value $Z(u)$ at each spatial location u so we use the distance map function from `spatstat` to calculate the distance from u to the nearest road. We then scale the covariate by 1000 so that the units are in kilometers.

Livelihood zones We grouped livelihood zones that were similar into 11 groups from the original 21 zones. A description of the groups is found in Appendix B. We converted the livelihood zones shapefile to an object of spatial class tesselation using the `spatstat` package. We then count the number of points in each region and estimate the intensity in each of the livelihood zones and rescale the covariate by 1000 so that the units are in kilometers to match the point pattern.

Power Plants First we rescale the power plant covariate by 1000 so that the units are in kilometers. Then we define a discrete binary variable that takes the value of 1 for locations within a 25 km radius from the power plant and 0 otherwise.

Population Density The population variable is the population density in which the number of people per pixel represents the number of people in a 100 x 100 m area. We reduced the resolution of the image to 256 x 256 pixels to improve the computational efficiency.

4.3.3 Diagnostics

To check whether our models are faithful to our dataset, we evaluate the models using several diagnostics including the AIC, lurking variable plots and Q-Q plots. The diagnostics are largely built from the residuals, which we obtain by subtracting the fitted mean from the observations (i.e. the data itself). Residuals capture important information about the discrepancy between the data and the models. We present the residuals in different ways (outlined in Table 4.2) in order to answer different questions. We can use a variant of the Akaike Information Criterion (AIC) to compare across the inter-

Table 4.2: Diagnostic tools for validating a fitted point process model. Modeled after Table 11.1 in Baddeley et al. (2015).

Target	Tool	Figure	`spatstat` Method
Fitted intensity	Relative intensity	4.2	`density.ppp`
	Smoothed residual plots	4.3	`diagnose.ppm`
Presence of covariate effect	Lurking variable plots	4.4	`lurking`
Independence	Q-Q plot of residuals	4.5	`qqplot.ppm`

action specifications. The best fitted model will have the lowest AIC. Lurking variable plots are helpful for detecting misspecifications in the trend component of a fitted model. Quantile-Quantile Diagrams (Q-Q plots) are graphs that compare the quantiles of two distributions, which help us check for misspecifications of the interaction component of a fitted model. Lastly, we use a smoothed version of the residuals which are shown in Figure 4.3 and are discussed in section 4.4.3. Smoothed residuals primarily reflect misspecification in the trend. The Q-Q plots and smoothed residuals are both based on simulations from the fitted model compared to the data. Two other diagnostics that are checks on overall fit (trend and interaction) are the inhomogeneous L-function and simply the distribution of the number of points from each model simulated point pattern.

4.3.4 OpenCellID Data

We use a publicly available source of cell tower locations from OpenCellID. OpenCellID is a community project that aims to geolocate cell towers around the world. The publicly accessible database includes information on cell tower locations and heights (Ulm et al., 2015). The cell tower locations are aggregated based on crowdsourced cellular signal measurements. The crowdsourced measurements are collected through individual records via smartphone apps such as Rf Signal Tracker, Tower Collector, and Keypad-Mapper-3 (OpenCellID, 2018a). These apps collect cell tower measurements while a user is mapping something else (OpenCellID, 2018b). For example, the app Keypad-

Mapper-3 records cell tower data while one maps house numbers and Points of Interests to contribute to OpenStreetMap[1]. The measurements can also be collected via hardware such as Raspberry Pi[2]. A user can access the data via the API and pull raw data which consists of the cell tower's latitude and longitude, the number of samples collected, and the day and time the tower was first detected.

We use the OpenCellID data for the country of Zambia which has Mobile Country Code (MCC) 645. Zambia had 30,872 measurement samples when the data were downloaded on October 3 2017. We cleaned the data by removing duplicate points and points that fell outside of Zambia's borders. Zambia has three mobile phone networks which were indicated by the Mobile Network Code (MNC) in the OpenCellID dataset: MNC 1 is Airtel (operated by Bharti Airtel); MNC 2 is MTN (operated by MTN Group); and MNC 3 is Zamtel (operated by Zambia Telecommunications Company Ltd.) We use the points from Zamtel (MNC 3) because it had the least number of cell towers and thus reduced computing time for fitting the models. Out of the three mobile phone networks in the country, Zamtel is the only government-owned telecommunication service provider.

4.3.5 Errors associated with crowdsourced cell tower data

There are a few known errors associated with the cell tower data, which was the best available at the time of analysis. These errors are related to the crowdsourced nature of the data which relies on individual measurements as well as the fact that multiple antennas are located on a single tower. Crowdsourced cell phone tower estimates are collected via the following process. First, a phone receives the radio signal from a GSM

[1] `http://keypad-mapper.org/wiki/Main_Page`

[2] In an email communication with an OpenCellID staff, it was confirmed that the database is completely contributed by the community. OpenCellID does not collect this information from governments or MNCs. Rather, volunteers use hardware devices or apps to submit measurement data. Thus, cells found in rural areas are from a user's device that physically visited the area where the cell was found and those measurements are uploaded to the database.

base transreciever station, which is located at the cell tower, and the GPS position of the cell phone and a unique ID pertaining to the base station is recorded. Often times measurements from the same base station are received at different places and therefore an average GPS position is stored in the OpenCellID database. When many measurements from different GPS positions are made from a single base station, then the precision of the averaged location is improved. This process of data collection leads to a few possible errors. In remote and less urban locations where fewer individuals have smartphones that record this kind of information, there are fewer measurements of base stations in general (if they are recorded at all) and also the averaged locational precision is decreased.

Additionally, there can be discrepancies between the actual position of a tower and that position reported by OpenCellID for two reasons. It is often the case that multiple antennas, i.e. "cells" are mounted onto a cell tower. This is because the antennas access different networks such as Long-Term Evolution (LTE) wireless broadband communication LTE, Universal Mobile Telecommunications System (UMTS), etc. In Germany, for example, the company Vodafone reports less than 40,000 cell towers, however, OpenCellID has more than 290,000 Vodafone cell ID records as of August 2014. Thus, each cell tower has more than seven antennas (i.e. cells) (OpenCellID, 2018b). We discuss this source of error in greater detail in terms of the results pertaining to the inter-point interaction in section 4.5.2.

Another reason why there are discrepancies between OpenCellID's location of towers and the actual location is that the crowdsourced measurements used to approximate the cell's location are not equally distributed around the tower. In other words, the sample collection process is uncontrolled and random. In cities, mobile phones may pass a cell tower from every side, i.e. one may be able to pass by a cell tower from all angles whether that is on a highway, walking on a street, or sitting at one's home. However, there can also be the case where a cell tower is on a hill with a road on just one side. Therefore, the measurement can only come from one angle which is a problem because cells on a tower emit signals that are often less than 360 degrees around the tower. In this case where the measurements are not equally distributed around the cell tower, the averaged GPS positions from these measurements is not going to be accurate. Overall, while the OpenCellID data are imperfect, they were the best available at the time of analysis and permitted a country-wide study of telecommunications access.

4.4 Results

4.4.1 Point pattern description

We use the cell tower locations of cell company Zamtel (MNC 3), which is a planar point pattern with 4,788 points. The window that contains the point patterns is a polygon of the country boundary of Zambia with units in kilometers. We show a kernel smoothed representation of the density of points in Figure 4.2. We see that there is a spatially varying intensity since there are more points in certain regions of the country than others and thus there is an inhomogeneous density. We see that cell towers become less abundant moving from the South to the North. There are two regions of high cell tower intensity which are the capital, Lusaka, and the Copperbelt region situated between northern Zambia and southern Democratic Republic of the Congo (DRC). The overall intensity of the point pattern is 0.00636 points per square kilometer.

4.4.2 Spatial process models: Model estimation

We consider the point process from one company (Zamtel) and we fit models with three covariates: roads, livelihood zones, and power plants. For the interaction component, we consider two forms—Geyer Saturation and Piecewise Constant—and an inho-

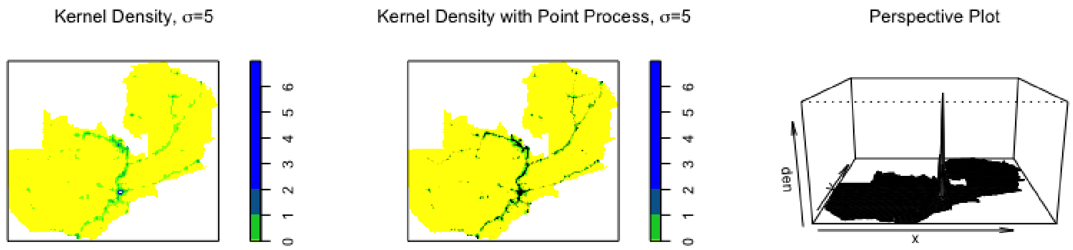

Figure 4.2: Density plots

mogeneous Poisson process. Our previous work shows that the point process is likely not an inhibition process because the models using Strauss Hard Core, Soft Core, and Diggle Gratton resulted in a singular matrix.

Table 4.1 shows trend components and the estimated interaction effects from the three functional forms, which are also depicted graphically in Figure 4.7. The significance of parameters are in comparison to hypothesis tests against the nulls for the trend (H_0 : $\eta = 0$) and interaction ($H_0 : \psi = 0$). In all of the models, the majority of covariates have a significant trend effect. The only covariate that is not significant is livelihood zones 3 and 7+ (across all models); livelihood zone 2 (PC model, only); and livelihood zone 4+ (GS model, only). The roads have slightly negative trend effects while most of the livelihood zones and power plants have positive trend effects. Livelihood zone 2, and to a lesser extent zone 7, are different compared to the other livelihood zones such that they have a negative trend effect across some or all of the models. For the roads covariate, moving away from the road decreases the conditional intensity. The largest effect sizes are livelihood zones 6+ and 8+ with a strongly positive trend component. The most negative effect size is seen in livelihood zone 2.

We also look at the consistency of the parameters across the different interaction functions. In terms of the sign of the estimate, we see that for each of the covariates the

Table 4.3: Gibbs model estimations for Zamtel point process with three environmental covariates: Roads, livelihood zones, and power plants. Standard errors are reported inside the parentheses. Asterisk denote significance at 1, 5, and 10 percent level. Livelihood zones denoted with "+" are groups of multiple livelihood zones.

Variable	None (Poisson)			Geyer-Saturation			Piecewise Constant		
Intercept	-6.66	***	(0.21)	-6.44	***	(0.17)	-6.53	***	(0.27)
Roads	-0.57	***	(0.01)	-0.43	***	(0.01)	-0.46	***	(0.01)
LZ 2	-0.92	*	(0.46)	-1.95	**	(0.72)	-1.21		(0.63)
LZ 3	0.54		(0.30)	0.34		(0.29)	0.28		(0.35)
LZ 4+	0.80	**	(0.25)	0.19		(0.27)	0.72	*	(0.32)
LZ 5	1.43	***	(0.27)	0.96	***	(0.26)	1.12	***	(0.32)
LZ 6+	2.89	***	(0.21)	1.71	***	(0.17)	3.05	***	(0.26)
LZ 7+	0.02		(0.31)	-0.09		(0.31)	0.07		(0.35)
LZ 8+	3.10	***	(0.21)	2.57	***	(0.17)	2.56	***	(0.26)
LZ 10	1.11	***	(0.25)	0.97	***	(0.26)	0.79	*	(0.33)
LZ 16+	0.91	***	(0.21)	1.02	***	(0.18)	1.05	***	(0.27)
LZ 18+	1.17	***	(0.24)	1.01	***	(0.21)	1.11	***	(0.29)
LZ 19	1.11	***	(0.22)	1.22	***	(0.19)	1.20	***	(0.28)
Power Plants	2.58	***	(0.04)	1.65	***	(0.04)	1.96	***	(0.04)
Interaction				0.02	***	(2e-4)			
Radius					4				
Saturation					181				
Radius Jump Points							0.5, 1.5, 2.5, 3.5, 4.5, 5.5		
AIC		32,337			20,341			22,050	

Table 4.4: Gibbs model estimations for Zamtel point process. Standard errors are reported inside the parentheses. Asterisk denote significance at 1, 5, and 10 percent level. Radius and saturation for Geyer Saturation are 4 and 8, respectively.

Variable	None (Poisson)			Geyer Saturation		
Intercept	-6.27	***	(0.07)	-6.52	***	(0.10)
Primary Roads	2.57	***	(0.06)	1.64	***	(0.07)
Secondary Roads	0.92	***	(0.05)	0.78	**	(0.07)
Log Population	0.63	***	(0.01)	0.57	***	(0.01)
Mining and Commercial	1.26	***	(0.05)	1.00	***	(0.05)
Most Food Insecure	-0.25		(0.14)	0.20	*	(0.10)
Interaction				0.21	***	(0.02)
AIC	17,223			13,783		

sign is consistent across the interaction functions except for livelihood zone 7+ in the GS model. In terms of consistencies in the significance level, the covariates are generally consistent across the interaction functions except in a few cases. We find that livelihood zone 10 is not consistent in Piecewise Constant since it is only significant at the 10 percent level. Additionally, livelihood zones 2 and 4+ have different significance levels across the interaction functions. We see that the covariates are fairly consistent in terms of the magnitude of the effect size across the interactions. The livelihood zones with the most disagreement in magnitude are zones 2 and 4+.

In summary, we find that the roads and power plants are consistent across effect size sign and magnitude. Certain livelihood zones are less consistent across the interactions, which is described further in section 4.5. We find that the Geyer-Saturation interaction has the lowest AIC and thus best overall fit. Additionally, all of the models with interaction have better overall fits than the inhomogeneous Poisson model.

Further details on livelihood zones

We proceed with a description of livelihood zones and explanations for differences in the point pattern's conditional intensity between zones. We refer to the grouped livelihood zone descriptions and original descriptions (FEWS NET, 2014) shown in Table B.1. The livelihood zones with the strongest and positive estimates across all of the models are "mining and commercial" (livelihood zone 6+) and "urban" zones (livelihood zone 8+). These areas of strong associations with the conditional intensity of cell tower locations are located in the central part of the country around Lusaka and spread north toward the Copperbelt region. These zones, particularly the mining and commercial zones, have low to very low risk of food insecurity. The only zone with moderate food security risk is zone 8 within 8+ which is the most agriculturally focused zone of the urban zones and experiences climate shocks approximately every 2-3 years, which affect agricultural production (FEWS NET, 2014).

The livelihood zones related to mining and commercial activities are 11, 6 and 13. The Mkushi, Chisamba, and Mpongwe commercial farming block (zone 11) is characterized by medium- to large-scale commercial farms. The area is a high-potential agriculture region with 800-1000 mm of rainfall per year and fertile soils (FEWS NET, 2014). The road infrastructure in this zone is good which facilitates movement of goods within and out of this zone. This zone is a net exporter of agricultural produce in Zambia. The Solwezi and Kasempa Mining Labor and Agricultural zone (zone 6) is characterized by copper mining, which is a recent shift in livelihood in this area (FEWS NET, 2014). The zone's proximity to the border of the DRC, decent roads, and communication infrastructure lead to good market access. Lastly, the Copperbelt Labor and Trade zone (zone 13) is a highly-urbanized highland area in which mining, timber, trade, and manufacturing are the main industries. The mineral resources that are mined in this area are copper, cobalt, amethyst, emerald and malachite with copper mining being the main economic activity. Agriculture is limited in this region compared to most of the other zones. This zone has urban centers and infrastructure that leverage it as a center for trade (FEWS NET, 2014).

The areas grouped as "urban" areas relate similarly to the cell tower point process with strong positive and significant effects. The southern plateau cattle, maize, and tobacco zone (zone 9) has a low risk of food insecurity and receives 800-1000 mm of rainfall per year. The zone has generally good infrastructure and several big towns are located within the zone that provide markets (FEWS NET, 2014). The central copperbelt maize, cassava and sweet potato zone (zone 12) also has a low risk of food insecurity. Rainfed and manual crop production is the main economy within the zone and the zone has both a good road and rail network which connects growers to markets in three main areas: Lusaka, Copperbelt and Solwezi (FEWS NET, 2014). The last zone in this group, the commercial rail line maize, livestock, and cotton (zone 8), is a little different in that it has moderate food security risk and is more densely populated at 19 people per square kilometer. Rainfed and irrigated agriculture are the main sources of livelihood while commercial and mechanized farms are scattered around the zone as well. The access to markets in this zone is good due to physical infrastructure which facilitates trade (FEWS NET, 2014). Although there is variability in communications access within livelihood zones, the cell tower access on average is better these urban zones compared to other livelihood zones.

The livelihood zones with weak and even negative estimates with cell towers are spread in a few areas, specifically the southwestern side of Zambia along the border of Namibia and Angola (livelihood zone 2); the western and northwestern area bordering Angola and DRC (livelihood zone 3); and two fishing areas: the north eastern and central areas of livelihood zones 7 and 21. The western and north-western cassava, maize, and cattle zone (zone 3) is considered one of the most at risk for food insecurity (FEWS NET, 2014). It is a region with low agricultural production potential due to poor sandy soils and flooding which experiences a high frequency of droughts. The zone is in a highland area with a longer growing season of 140-160 days. Individuals in these areas have less access to major markets due to poor communication and road infrastructure. The southwestern cereal, livestock and timber zone (zone 2) is characterized by moderate food security risk and is located in agro-ecological region 1 in which less than 800 mm of rainfall occurs and the main growing season is short (between 90 and 120 days). The road infrastructure in this area is generally good and leads to satisfactory access to major markets in Livingstone (FEWS NET, 2014).

4.4.3 Diagnostics for the trend and interaction components

The Gibbs point process models have both a coherent inferential framework and are parametric, and thus the interaction and trend can be evaluated using various specification tests and diagnostics (Sweeney and Arabadjis, 2021). We also use the models to simulate the point pattern process and compare the simulated pattern against the observed pattern. The smooth residual plot helps us draw attention to regions in Zambia where there is the largest difference between the model and the data. Figure 4.3 shows the kernel-smoothed raw pseudo-residuals and where in Zambia the models are over- and under-fitting the spatial trend. For both models (Geyer Saturation on the left and Piecewise on the right) the residuals are close to 0. The range of values is between -0.003 and 0.004 for both subplots. The figures show that the region nearby Lusaka has slightly more positive residuals (between 0.001 and 0.004) whereas the rest of the country is zero or slightly negative towards the Southern Province of Zambia (between 0 and -0.003). The areas in which the residuals are slightly positive indicates that there are more points observed than were predicted. The areas with more negative residuals suggest over-prediction. Therefore we find that the model is not fitting well around the capital, Lusaka, which is discussed further in section 4.5.1.

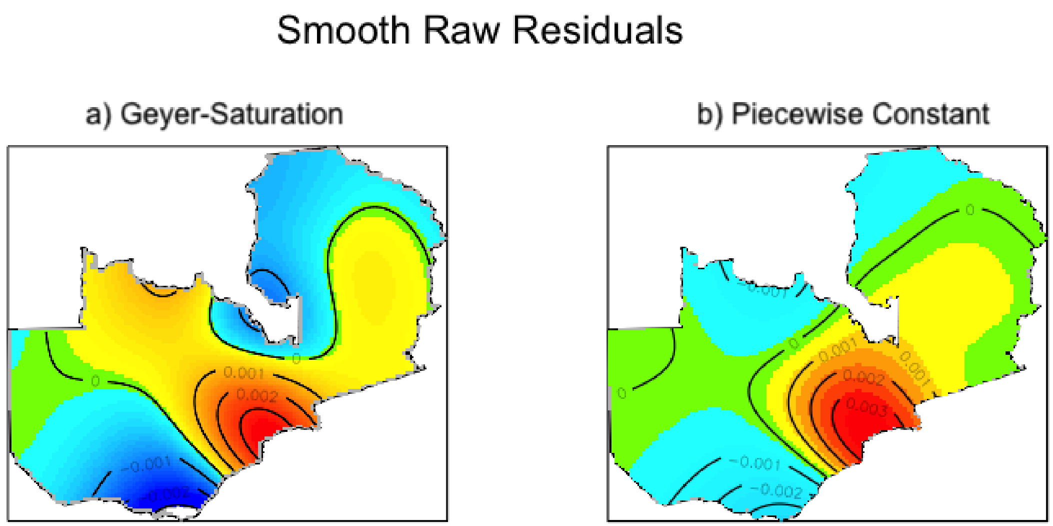

Figure 4.3: Maps of smoothed raw pseudo-residual fields: a) Geyer-Saturation and b) Piecewise constant.

Lurking variable plots

Lurking variable plots show us where the model over- or under-predicts points with respect to an explanatory variable. Figure 4.4 shows the roads covariate on the x-axis and the cumulative raw residuals on the y-axis. The lurking variable plot starts and ends at the cumulative raw residual of 0 because the residuals in this case are the observed minus the predicted values. The cumulative residuals and thus the total residual for the entire plot is equal to 0. The gray bands show the 5% significance bands for the cumulative residual. The roads lurking variable plot shows us that the error bands follow the shape of the residuals. Overall, we are under-predicting until ca. 15 km. There is also a significant spike in under-prediction of points occurring at very short distances. We discuss this spike in residuals at short distances in the discussion section 4.5. Overall this plot shows us that roads are an adequate predictor of the point processes' intensity

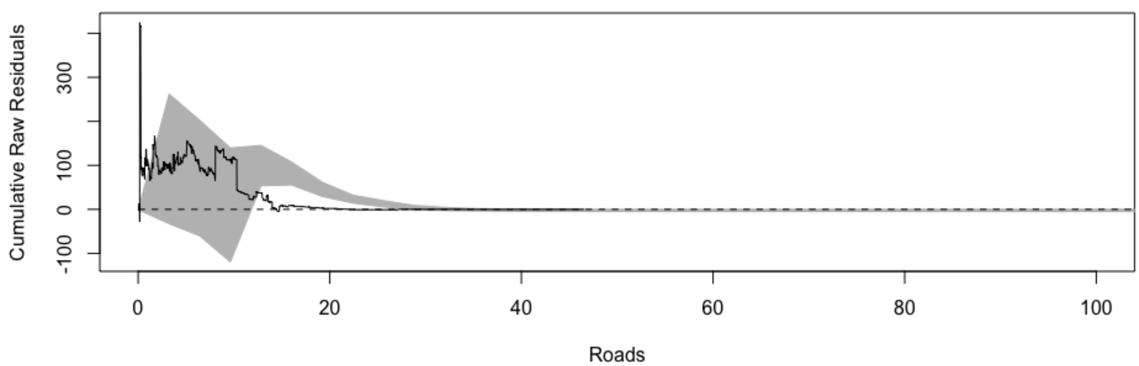

Figure 4.4: Lurking variable plots using the Geyer-Saturation interaction.

despite some under-prediction of points at small spatial scales.

Q-Q plots

We validate the interaction component of the model with the summary interaction function $L(r)$ and Q-Q plots of the pseudoresiduals. For two of our three models (Geyer-Saturation and inhomogeneous Poisson) we can fit Q-Q plots shown in Figure 4.5. For the PC interaction, the model is not valid and therefore we cannot generate a Q-Q plot. We see a slight improvement in the Geyer-Saturation Q-Q plot in comparison to the inhomogeneous Poisson plot. The confidence bands are very narrow in both plots due to the large number of data points. If there was perfect agreement between the point process data and the simulated point process we would see the points fall along the diagonal line. With the inhomogeneous Poisson Q-Q plot, we find a very poor relationship between the observed data and the Poisson process, which is indicated by the almost vertical nature of the points rather than falling along the diagonal. Thus, the uniform Poisson model is inappropriate for the cell tower point process data. The Q-Q plot for the inhomogeneous Poisson model shows us the influence of the trend without

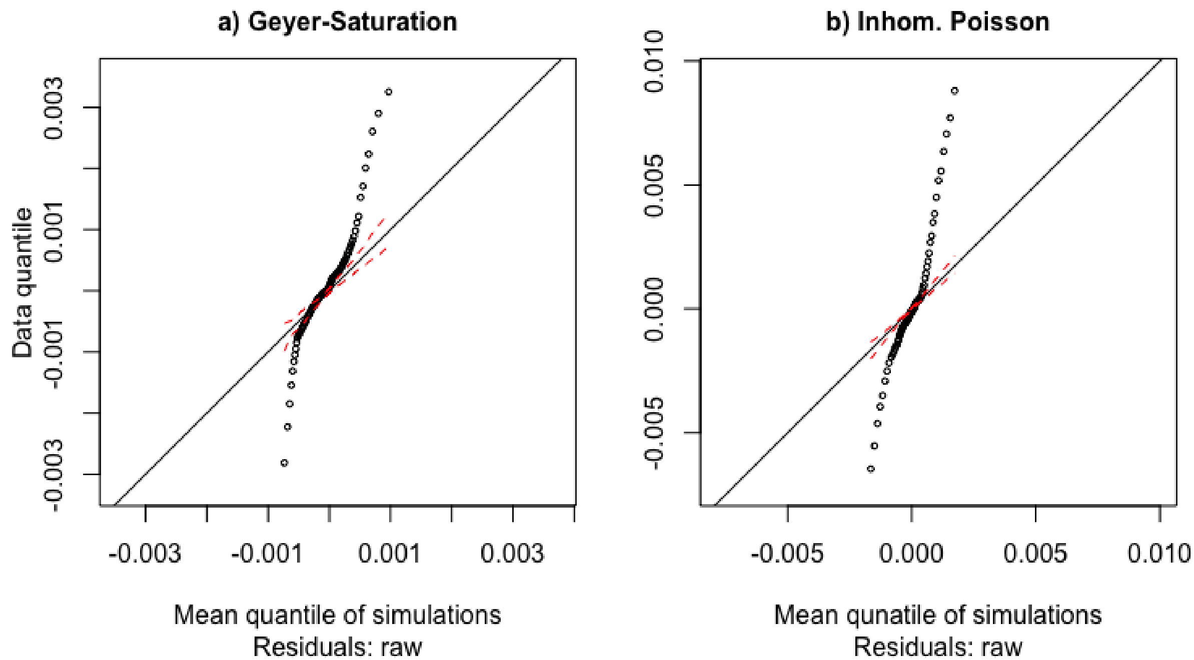

Figure 4.5: QQ-plots for a) Geyer Saturation interaction and b) inhomogeneous Poisson.)

any consideration of the clustering process or any other inter-point interaction. The Q-Q plot for Geyer-Saturation indicates that a clustering process is perhaps appropriate for the point process.

Inhomogeneous L-function

We use the summary interaction function $L(r) - r$ to diagnose the interpoint interaction as shown in Figure 4.6. We use the fitted model with the Geyer-Saturation interaction to run simulations of the point pattern and we select the minimum and maximum L-function values to construct uncertainty envelopes. The L-function is entirely within the uncertainty envelopes which indicates a good fit. This is because the Geyer-Saturation captures both the trend and interaction for the point process.

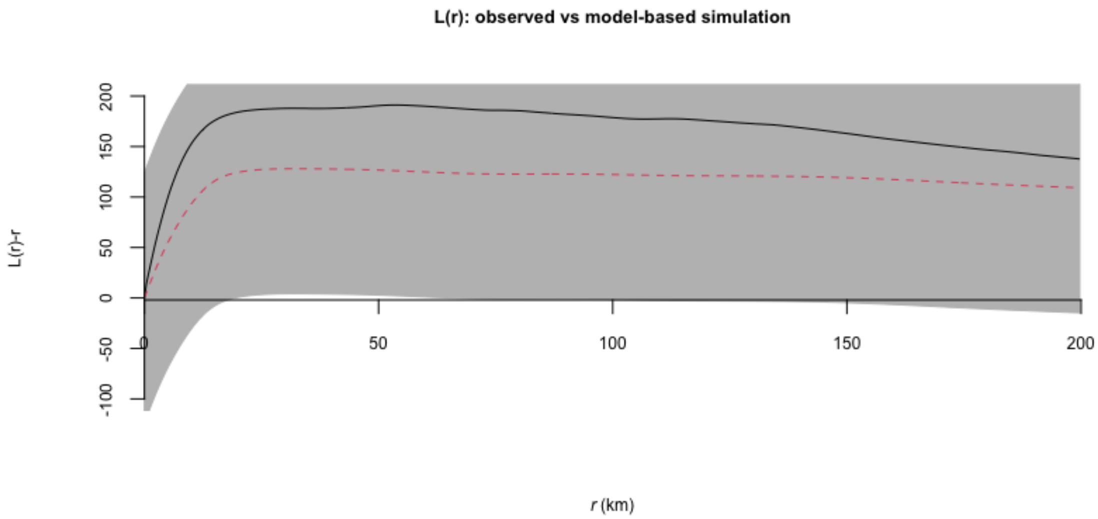

Figure 4.6: L-function of fitted Geyer-Saturation model

Fitted interaction functions

We plot the fitted interaction function for the Geyer-Saturation and Piecewise constant models in Figure 4.7. These visuals provide us with the fitted interpoint interaction. The y-axis is the function $h(r)$ which is the pairwise interaction between points for each of the models. Both of the plots show clustering rather than inhibition since values of $h(r)$ above 0 indicate clustering and values below 0 indicate inhibition. For GS interaction, clustering ceases at 4 km at which point there is no effect. The saturation for the GS interaction is 181 (Table 4.3), which further indicates clustering. For the PC interaction, there is clustering until approximately 1 km and then the interaction alternates between clustering and no effect between 1 and 7 km.

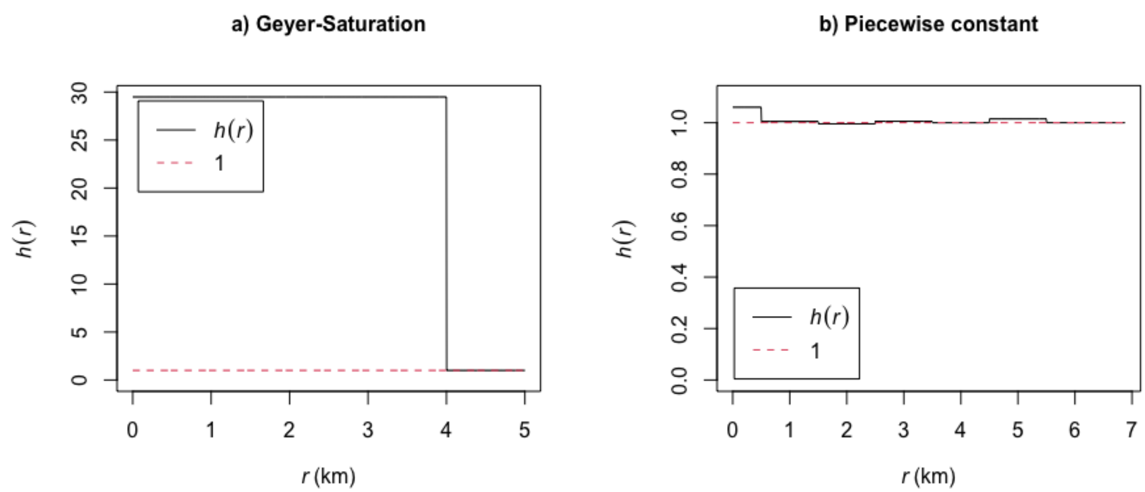

Figure 4.7: Fitted interaction patterns

Does the number of points that we simulate out of the point process roughly match the original number of points?

We find that the inhomogeneous Poisson, Geyer Saturation, and Piecewise Constant models under predict the number of points in Table 4.5. The models that are closest to the actual number of points (4,788) are the inhomogeneous Poisson model when only a single covariate is included, specifically livelihood zones or power plants. The Geyer Saturation model does not provide as close to the correct number of points as inhomogeneous Poisson. When single covariates are used, the GS model produces less than half of the actual number of points. These results indicate that improvements to model fitting with respect to GS is possible, and we are working on making improvements to the model.

4.5 Discussion

Table 4.5: Number of points simulated out of point process with 4,788 points

Covariates	Interaction	Number of Points	Ngrid
Roads, LHZ, PP	Inhomogeneous Poisson	3891	1000
	Geyer Saturation	2681	1000
	Piecewise Constant	-	
Livelihood zones	Inhomogeneous Poisson	4850	1000
	Geyer Saturation	2031	1000
	Piecewise Constant	-	
Roads	Inhomogeneous Poisson	4053	1000
	Geyer Saturation	1972	1000
	Piecewise Constant	-	
Power plants	Inhomogeneous Poisson	4793	1000
	Geyer Saturation	2038	1000
	Piecewise Constant	-	

4.5.1 How do socio-environmental variables relate to the location of cell towers?

Overall, we see strong relationships between roads, power plants, and certain livelihood zones and the intensity of cells. The livelihood zones that most strongly predict cell tower location are found along the railroad track between Livingstone and the Copperbelt region. These livelihood zones are characterized by copper mining, commercial farming, and urban economies. Livelihood zones that have weaker relationships to the conditional intensity of cell towers include more food insecure regions and those least dominated by mining and commercial activity. This is, however, a nuanced result because there is variability in livelihood and access to telecommunications within zones. For example, in livelihood zone 8+ which extends from Livingstone to Lusaka there are commercial activities taking place as well as smallholder agriculture. Smallholder agriculture occurs throughout Zambia, however, communities that are located further away from roads will inevitably have reduced access to cell towers.

Investigating which covariates are most strongly related to power plants provide insights into Zambia's history of economic development and present day economic and social inequities. Access to electricity is at the root of economic development and social welfare. The location of cell towers, which depend on a grid connection, provide a case study example into the inequities in access to communications infrastructure, and more broadly, electrification in a country. By identifying what forces led to the placement of cell towers in particular locations we gain insight into a public welfare issue that has implications for universal access in electrification and mobile phones in Zambia.

Are cell towers being placed for the needs of rural communities?

Roads are the main covariate that explains cell tower location which is evident from Table 4.3 as well as visually in Fig. 4.2. The point process' relationship to roads is also linked to the model's behavior in urban and rural areas. The model fits more poorly in the urban centers of Zambia compared to other regions (Fig. 4.3). The reasons why we believe under-predicting is occurring around Lusaka is due to the high number of points clustered together in urban areas. This is due to two reasons: (1) The location of cell towers are actually the placement of antennas on cells and (2) there are more crowdsourced measurements in cities due to the higher population count. The model fits improve greatly in rural areas, which is likely due to the cell towers being placed very close to roads. This is discussed further in the next section.

Linkages between access to communications and power infrastructure in Zambia

We find that the development of communication services and infrastructure and energy production are inextricably linked. First, the use of ICTs and access to e-services are constrained by the availability of electricity. In other words, an individual needs electricity to be able to charge their mobile device and maintain its battery. Additionally, the communications network that connect mobile phones and the actual installation of the cell tower requires electrification. Therefore, regions of Zambia without gridded electricity access are unlikely to have cell towers. Only 4 percent of rural Zambia has power and thus most rural residents do not have access to electricity at their homes (Mudenda et al., 2013).

Our proxy for access to electricity is the location of power plants. We find that power plants have strong and positive associations with the conditional intensity of cell towers. The range of the estimate is between 1.6 and 2.6 across the models. Power plants are not evenly distributed across the country with the majority of power plants being located in livelihood zones 6+ and 8+. Of the 83 power plants, the majority are hydroelectric stations and a smaller sample is thermal.

It is no accident that hydroelectric stations, mining resources, and cell towers are co-located whereas areas without road networks and regions with subsistence farming are underserved. Zambia's National Grid infrastructure spans 2,241 km across the country with the main artery running between Livingstone and the Copperbelt region (Figure 13 in (World Bank, 2017). This pattern mimics the major highway between Livingstone and Copperbelt. In the south, the main hyropower stations are located at Kariba North and Kague dams and the grid connects to the north where the main load centers are located. These large hydro plants have dominated Zambia's power generation for decades (Bayliss and Pollen, 2021) and the load centers were selected to be located in the Copperbelt area due to mining activities.

Zambia's economy has a heavy reliance on mining and there is very little electrification in rural areas. There is a large divide in electricity access between urban and rural households. About 57% of the population lives in rural areas but only about 4% of these households have access to electricity whereas 67% of urban households do access electricity (World Bank, 2017; CSO, 2004). Of the poorest 20% of Zambians, less than

1 percent receive electricity subsidies compared to the richest 20% of the population who receive about 70% of the subsidies (Kabechani et al., 2017). Furthermore, recent droughts have led to a crisis in hydropower which are exacerbating some of the issues with how electricity is provided and consumed in Zambia (Bayliss and Pollen, 2021).

4.5.2 Do cell towers exhibit a cluster or inhibition interaction?

For the interaction component of the spatial point process, we originally hypothesized that cell towers would exhibit an inhibition interaction. This hypothesis was based on the notion that cell towers should not be co-located in order to maximize services to users spread throughout a large area. Rather, we found that clustering occurs at very small spatial scales. In the lurking variable plot of the roads covariate with the Geyer-Saturation interaction, Fig. 4.4, we see extremely high cumulative raw residuals at very small spatial scales. This indicates that we are under-predicting points very close to roads. Furthermore, in the maps of the smoothed raw residuals we found that the positive pseudoresidual fields are concentrated around the capital Lusaka and improve elsewhere in the country. This is because of the extremely high density of points in the capital. We are working on the next iteration of the model which improves the tendency for points to be clustered in the capital due to multiple antennas being placed on a single cell tower.

The tendency for cell towers to be clustered is also be an oddity of the crowdsourced method of data collection. The layout of how crowdsourced samples are collected differs in urban, peri-urban and rural areas (Fida and Marina, 2019). For example, pedestrians on sidewalks and passengers along streets and roads in cities will lead to data being collected that allows for all angles of a cell tower to be recorded. However, cell towers are sometimes not accessible from all directions, e.g. when they are placed on a hill, and thus the collection points are not as limited. The location of a tower, the cell tower's footprint, the population density, and the moving pattern of the individual with the cell phone will affect how the cell tower's location is triangulated (Fida and Marina, 2019). In conclusion, we do not find much evidence for inhibition. We ran models for the Strauss Hard Core, Soft-Core, and Diggle Gratton interaction functions and could not produce viable point patterns. In addition, the fitted interaction patterns for Geyer-Saturation and Piecewise constant, Fig. 4.7, indicate that there is either clustering or no effect rather than inhibition.

4.6 Concluding remarks

Our goal for this paper was to model the spatial pattern and inter-point interaction of cell towers in Zambia. Understanding access to communications infrastructure is important for a wide variety of applications including but not limited to early warning systems, food security, poverty mapping, and the diffusion of government and other services. We applied Gibbs models to the spatial point pattern of Zamtel cell towers. First, we investigated whether the intensity of cell towers depended on the spatial covariates. Second, we studied the spatial dependence between points after allowing for covariate effects on the cell tower intensity. We identified trends in cell tower locations which are primarily situated in Zambia's two urban areas: Lusaka and Copperbelt. With the spatial process models we gained insight into the role of various environmental covariates and found that distance to roads and location within an urban and/or industrial center is most closely related to the conditional intensity of cells. The fitted interaction functions indicated that cells exhibit clustering in which the deployment of a cell tower in one location gives rise to another one nearby. Upon analyzing the data through simulations and specification tests, we determined that the Geyer-Saturation interaction is the most likely function for representing the point pattern.

The use of mobile phones or internet is predicated on the availability of electricity. Zambia's energy sector is highly dependent on hydroelectric power which is susceptible to droughts. Thus, regular power outages occur that interrupt mobile phone services and Internet (Mudenda et al., 2013). Our results highlight the vulnerability of Zambia's communications and thus power infrastructure such that it is concentrated in just a few sectors and in a specific region in particular along the corridor between Livingstone and Copperbelt. With such low rural electrification rates in Zambia, the solution to improved telecommunications access for all Zambians is not as easy as simply installing power infrastructure or cell towers. Because individuals in villages are disconnected from roads, power, and telecommunications, it is not always appropriate to install IT equipment before providing other vital resources such as a community health clinic, for example (Mudenda et al., 2013). Thus, identifying priority areas for electrification needs to be based on community input.

Lastly, more of a focus on social equity and environmental sustainability is needed on studies of communication access and electrification. We have already highlighted the divide in telecommunications access, poverty, and electrification between rural and urban parts of Zambia. Given the landscape of increasing participation from Independent Power Producers (IPPs) and increased private investment in Zambia's infrastructure, one study concluded that the revenues made by investors' returns could easily pay for rural electrification (Bayliss and Pollen, 2021). The energy sector reforms that took place in the 1990s including in Zambia have not proven to have positive social outcomes. Rather, the private sector, intergovernmental agencies, and international institutions such as the World Bank need base their policies on local circumstances: what works in one country should not be taken as the standard or default to work in another. Further advancements to our understanding of the location of telecommunications access in the Global South should center issues of equity, the industrial and colonial history of the place of interest, and how persistent inequalities result in disparities in telecommunications seen today.

Chapter 5

Conclusion

How the world will continue to feed itself while ensuring equitability and sustainable use of natural resources is an open question. As the planet warms, weather and climate variability will increase (Thornton et al., 2011; IPCC, 2014). These impacts are already affecting agricultural systems globally through changing precipitation patterns, increasing temperatures, and increasing frequency of extreme events (Mbow et al., 2019). Climate change is set to continue disrupting systems with climate-related disasters impacting countries in the Global South disproportionally (Hansen et al., 2019). With current levels of anthropogenic greenhouse gas emissions, the deleterious impacts of climate change and climate variability are unlikely to improve in the short-term. Given this, proactive early warning systems are a critical point of research and point of international collaboration.

Digital tools can support agriculture in a number of ways. From food security, access to markets, and government and extension services, every facet of the food system can benefit from technology. Actions such as emergency assistance, early warning, and distribution of critical resources are needed to prevent famine, particularly through assistance that is scalable and remotely deployable. Mobile phones are one pathway that can be an important step towards providing assistance in multiple forms. Digital agricultural systems span multiple scales from farm-level tools and applications that can deliver services of direct benefit to farmers to ecosystem- and macro-level tools that create an enabling environment for broadening participation in digital services (Gates Foundation, 2020).

The gap in coverage and internet access is more important to study than ever as the world's reliance on internet, specifically via access through mobile phones, has never been stronger (Bahia and Suardi, 2019). The final year of research on this dissertation coincided with an unprecedented global pandemic. The impacts of COVID-19 led to an uptick in acute hunger in which over 34 million people faced emergency levels of hunger around the globe (WFP and FAO, 2021). The COVID-19 pandemic demonstrated the importance of mobile phones and internet information diffusion in particular for the majority of the world who accesses the internet through a mobile device (Bahia and Suardi, 2019). An estimated quarter of the population in sub-Saharan Africa live outside of mobile broadband coverage whereas the global average living outside of coverage is seven percent (Bahia and Suardi, 2019). Because the primary way that individuals access digital information is through mobile broadband, studying access has never been more important.

Finally, small-scale farmers in sub-Saharan Africa are often the first to feel the impacts of extreme weather. The issues of food security, climate change, health, land degradation and conflict are all interlinked and technology-based solutions are just one of many paths forward for reducing vulnerabilities associated with a changing climate. Small-scale producers, particularly women and those who are disconnected from road networks, farm marginal lands, and have weak connections to markets need to be given particular attention. With rapidly improving broadband coverage and mobile-phone usage across rural populations in sub-Saharan Africa, integrating technology with smallholder farmers' strategies for maximizing production while minimizing resources and land degradation should be an opportunity for advancing towards food security on the continent.

Appendix A

Chapter 3 Appendix

A.1 Uncertainties in maximum yields

We use seed company-provided estimates of hybrid maize yields to set maximum values of yields by maturity period. These values should be considered as the "potential yields" for a given variety because they are usually attained from optimal growing conditions and from the possible addition of agricultural inputs and management (fertilizer, weeding, etc.) that does not necessarily reflect small-scale farming conditions (Blekking et al., 2021). How to ascertain whether seed company-provided yields can actually be attained is through cultivar trials, which are done through local research stations that play an important role in certifying seeds and verifying seed company reported yields. Blekking et al. (2021) compared yields between maize companies, Zambia's Seed Control and Certification Institute (SCCI) and from smallholder farms for five varieties in Zambia during the 2011/12 growing season. They found that the seed developer stated potential yields, average yields verified through SCCI cultivar trials, and farmer yields differed vastly. Farmer yields were just a third or less than the company stated yields.

For the hybrid varieties of interest in our study, the range of maximum yields was roughly 1.5 to 5 t/ha as shown in Figure 3.3. However, our model simulations which were subjected to stochastic rainfall resulted in lower average yields with averages between 1.18 and 1.45 for early to late maturing crops (Table 3.5). Previous work has shown that regional yields vary and were an average of 0.92 - 1.55 t/ha between 2000 and 2014 in Laikipia county (Davenport et al., 2019). Given this information, our simulated maize yields fall within regional maize yields. Despite the differences seen between seed company yields and realized yields (Blekking et al., 2021) our model results align with some of the regional yields typically seen for our study site. Future work could better tune the realized yields for a range of maturity periods to better simulate the typical yields found at small-scale farms.

A.2 Sensitivity analysis

We conducted a sensitivity analysis to investigate whether model parameters related to static and dynamic water stress and evaporation affect the model results. These four variables (r, k, q_e, and q_{stress}) are interlinked and difficult to measure directly. We find that yields are particularly sensitive to the value of k in affecting the range of yields, median yield value, and likelihood of crop failure. The value of r also changes the range of yields but has a limited affect on the median yield. We settled on $r = 0.5$ as our value because Porporato et al. (2001) selected this response between stress and yield to be less than linear (ie. the square root of the term). The k parameter has the opposite effect where smaller values of k lead to lower yields whereas higher values lead to higher yields. The k parameter is the portion of the season that the crop can tolerate stress before it fails. Both of these values are interlinked with the calculations of dynamic stress and therefore end of season yields, and are tied to the crop's physiology. Our results show that these values impact the model results and therefore more data is needed on how intra-

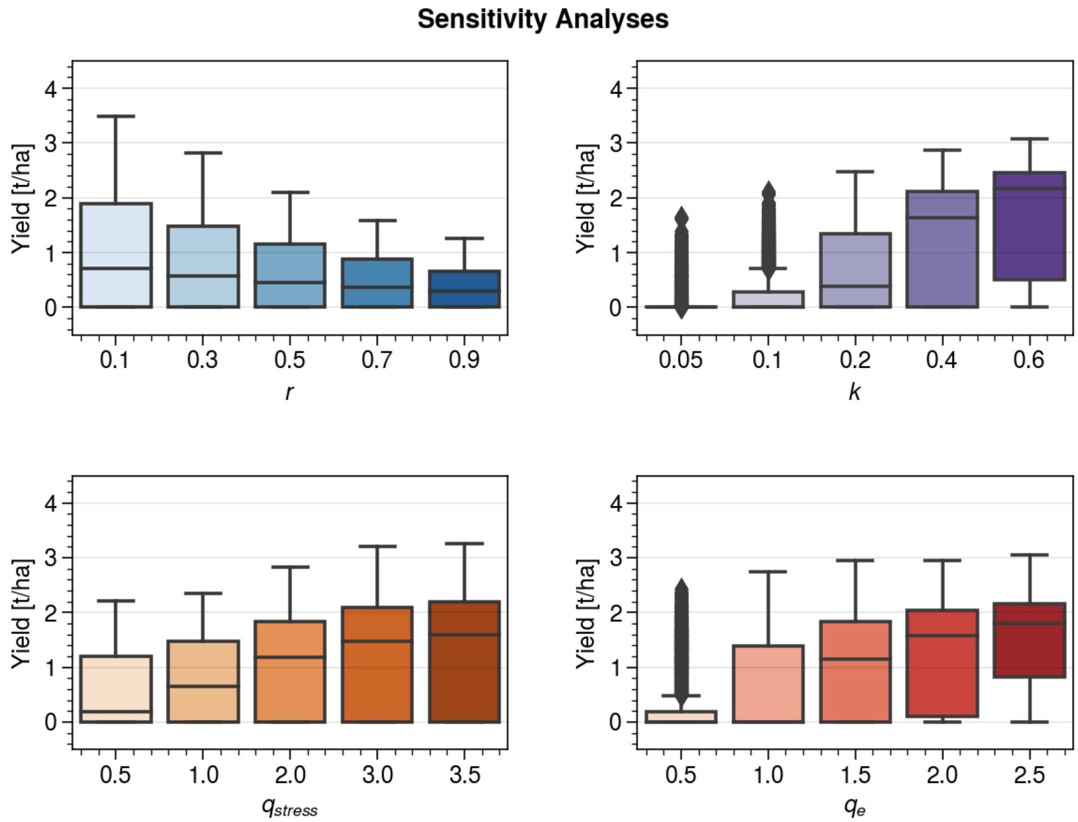

Figure A.1: Effect of different values of r, k, q_{stress} and q_e on yields.

seasonal water availability affects the duration and severity of stress and therefore crop outcomes. These values are both difficult to resolve among variety types and are difficult to disentangle since they are both tied to the crop's physiology. Thus very little data exists on the effects of intraseasonal crop stress on yield outcomes. Crop modeling studies often only work with seasonal rainfall totals or average weather rather than within season rainfall variability. Increased attention needs to be paid to these parameters which would be improved through field trials and real-time estimates of crop growth under varying meteorological conditions.

A.3 PDF of dynamic stress

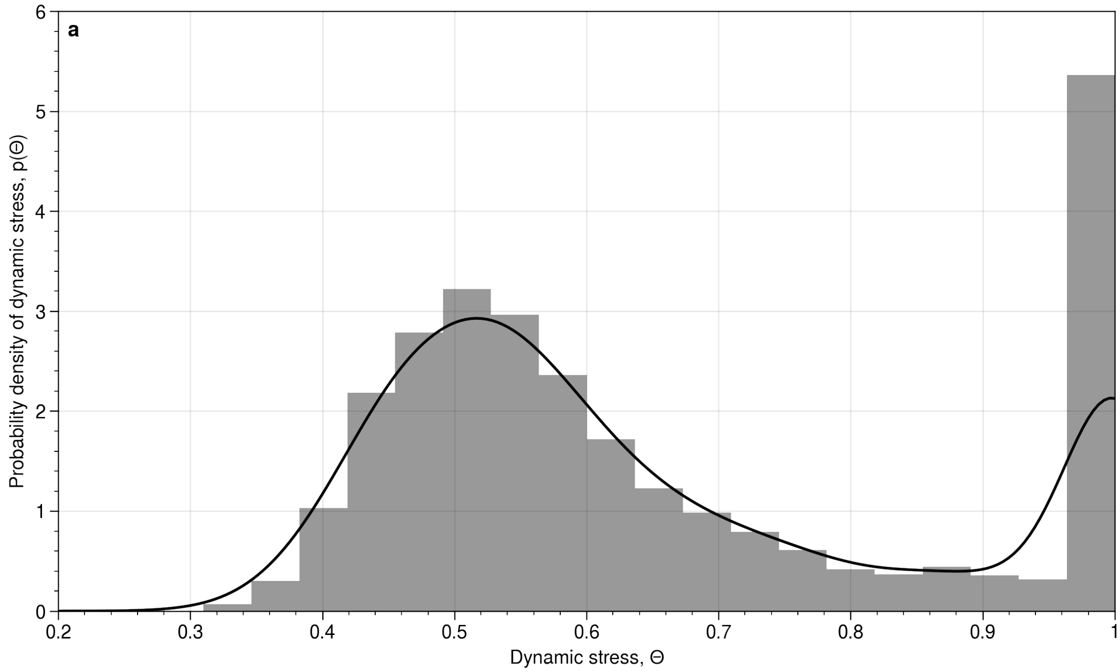

Figure A.2: Probability distribution function of dynamic water stress for 10,000 simulations. Model parameters used are the same as those in Figure 3.6.

A.4 Rainfall-yield relationships for varieties

There is a wide range of yields and therefore farmer outcomes associated with certain bands of rainfall as shown in Figure A.3. In the early maturing varieties, we are concerned about the divergence of yield outcomes within the band of 100-200 mm of seasonal rainfall; 150-300 mm of rainfall for medium maturing, and 200-400 mm of rainfall for late maturing. For these bands of rainfall, we find that a wide range of outcomes are possible for farmers such that one farmer (i.e. one model realization) that experiences 175 mm of seasonal rainfall for an early-maturing variety can realize about 75 percent of the crop's maximum yield while another farmer that receives 175 mm of rainfall experiences total crop failure. We find this relationship across all of the varieties categories in which some realizations get no yield or very little yields while others get 50-75 percent of their maximum yield for the same seasonal rainfall total. The only difference between these two outcomes (no yield vs. decent yield) is the timing of rainfall because the total rainfall is the same.

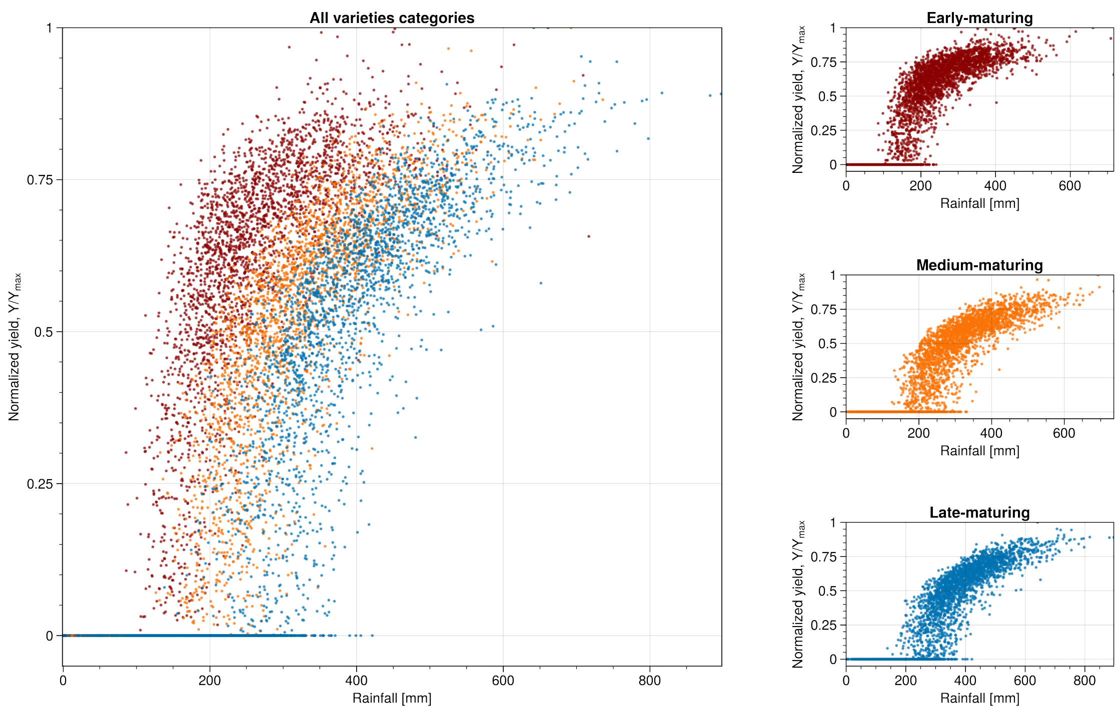

Figure A.3: Relationships between seasonal rainfall and normalized yields for three categories of varieties: Early maturing (red), Medium maturing (orange), and Late maturing (blue).

Appendix B

Chapter 4 Appendix

B.1 Livelihood zones groups

There are twenty one livelihood zones per FEWS NET (2014). We reduced the number of livelihood zones to eleven in order to simplify the description of the results. To do this, we first ran an inhomogeneous Poisson model with roads, power plants, and all 21 livelihood zones. We made an initial grouping of livelihood zones based on covariates that had similar coefficient values in the model estimates table. Second, we altered the groupings based on which livelihood zones were similar in their qualitative description as well as those which were geographically nearby each other. We then used 11 categories of livelihood zones which are found in Table B.1.

Table B.1: Grouping of 21 original livelihood zones into 11 categories. The original labels and categories are from FEWS NET (2014) and our new labels are categories shown as the first two columns of the table

New Grouping	New Category	Original Label	Original Category
Baseline		LZM01	Zambezi Plain Rice, Livestock, and Fishing
LZ 2		LZM02	Southwestern Cereal, Livestock, and Timber
LZ 3		LZM03	Western and Northwestern Cassava, Maize, and Cattle
LZ 4+	Tobacco	LZM04	Kaoma Smallholder Food Crop and Tobacco
		LZM17	Eastern Plateau Maize, Groundnut, Tobacco, and Trade
LZ 5		LZM05	Northwestern Timber and Honey
LZ 6+	Mining and commercial	LZM06	Solwezi and Kasempa Mining Labor and Agriculture
		LZM11	Mkushi, Chisamba, and Mpongwe Commercial Farming Block
		LZM13	Copperbelt Labor and Trade
LZ 7+	Fishing	LZM07	Kafue Plain Maize, Cattle, and Fishing
		LZM21	Mweru, Bangweulu, and Tanganyika Fisheries
LZ 8+	Urban	LZM08	Commercial Rail Line Maize, Livestock, and Cotton
		LZM09	Southern Plateau Cattle, Maize, and Tobacco
		LZM12	Central Copperbelt Maize, Cassava, and Sweet Potato
LZ 10		LZM10	Zambezi Valley Agro-fisheries
LZ 16+	Smallholder: Groundnut	LZM16	Eastern Plateau Maize, Cotton, and Groundnut
		LZM20	Muchinga, Northern and Luapula Cassava, Groundnut and Mill
LZ 18+	Luangwa Valley	LZM14	Mkushi and Serenje Maize, Sweet Potato, and Horticulture
		LZM15	Luangwa Valley Informal Mining, Fishing, and Hunting
		LZM18	Luangwa Valley Subsistence Farming, Hunting, and Tourism
LZ 19		LZM19	Northern Border Maize, Beans, Livestock, and Trade

Bibliography

Adger, W. N. (2006). Vulnerability. *Global environmental change*, 16(3):268–281.

Allen, R., Pereira, L., Raes, D., and Smith, M. (1998). Chapter 6: ETc-Single crop coefficient (KC). *Crop evapotranspiration—guidelines for computing crop water requirements—FAO irrigation and drainage paper 56*.

Appelbaum, S. H. (1990). Computerphobia: Training managers to reduce the fears and love the machines. *Industrial and Commercial Training*, 22(6).

Apurv, T., Sivapalan, M., and Cai, X. (2017). Understanding the role of climate characteristics in drought propagation. *Water Resources Research*, 53(11):9304–9329.

Baddeley, A., Rubak, E., and Turner, R. (2015). *Spatial point patterns: Methodology and applications with R*. Chapman and Hall/CRC.

Baddeley, A. and Turner, R. (2005). spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software*, 12(6):1–42.

Bahia, K. and Suardi, S. (2019). The state of mobile internet connectivity 2019. *GSMA Connected Society: London*.

Bapna, M., Brandon, C., Chan, C., Patwardhan, A., and Dickson, B. (2019). Adapt now:

A global call for leadership on climate resilience. Technical report, Global Commission on Adaptation, Rotterdam, Netherlands.

Barron, J., Rockström, J., Gichuki, F., and Hatibu, N. (2003). Dry spell analysis and maize yields for two semi-arid locations in east Africa. *Agricultural and forest meteorology*, 117(1-2):23–37.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Baudena, M., Boni, G., Ferraris, L., von Hardenberg, J., and Provenzale, A. (2007). Vegetation response to rainfall intermittency in drylands: Results from a simple ecohydrological box model. *Adv. Water Resour.*, 30(5):1320–1328.

Baumüller, H. (2013). Mobile technology trends and their potential for agricultural development. Technical report, University of Bonn, Center for Development Research (ZEF).

Baumüller, H. (2015). Assessing the role of mobile phones in offering price information and market linkages: The case of M-Farm in Kenya. *The Electronic Journal of Information Systems in Developing Countries*, 68(1):1–16.

Baumüller, H. (2018). The Little We Know: An Exploratory Literature Review on the Utility of Mobile Phone-Enabled Services for Smallholder Farmers. *Journal of International Development*, 30(1):134–154.

Bayliss, K. and Pollen, G. (2021). The power paradigm in practice: A critical review of developments in the Zambian electricity sector. *World Development*, 140:105358.

Becker-Reshef, I., Justice, C., Barker, B., Humber, M., Rembold, F., Bonifacio, R., Zappacosta, M., Budde, M., Magadzire, T., Shitote, C., et al. (2020). Strengthening agricultural decisions in countries at risk of food insecurity: The GEOGLAM crop monitor for early warning. *Remote Sensing of Environment*, 237:111553.

Besag, J. (1977). Contribution to the Discussion on Dr. Ripley's paper. *Journal of the Royal Statistical Society*, B39:193–195.

Bharwani, S., Bithell, M., Downing, T. E., New, M., Washington, R., and Ziervogel, G. (2005). Multi-agent modelling of climate outlooks and food security on a community garden scheme in Limpopo, South Africa. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 360(1463):2183–2194.

Blekking, J., Waldman, K. B., and Evans, T. (2021). Hybrid-maize seed certification and smallholder adoption in Zambia. *Journal of Environmental Planning and Management*, 64(2):359–377.

Braimok, T. (2017). Exploring the opportunities and challenges of ICTs for women farmers in Kenya. Unpublished master's thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden.

Branca, G., McCarthy, N., Lipper, L., Jolejole, M. C., and Others (2011). Climate-smart agriculture: A synthesis of empirical evidence of food security and mitigation benefits from improved cropland management. *Mitigation of climate change in agriculture series*, 3:1–42.

Bryan, E., Deressa, T. T., Gbetibouo, G. A., and Ringler, C. (2009). Adaptation to climate change in Ethiopia and South Africa: Options and constraints. *Environmental science & policy*, 12(4):413–426.

Burrell, J. and Oreglia, E. (2015). The myth of market price information: Mobile phones and the application of economic knowledge in ICTD. *Economy and Society*, 44(2):271–292.

Caine, A., Dorward, P., Clarkson, G., Evans, N., Canales, C., and Stern, D. (2015). Review of mobile applications that involve the use of weather and climate information: Their use and potential for smallholder farmers. CCAFS Working Paper no. 150. Technical report, CGIAR Reserch Program on Climate Change, Agriculture and Food Security (CCAFS). Available online at: www.ccafs.cgiar.org, Copenhagen, Denmark.

Camacho, A. and Conover, E. (2011). Manipulation of social program eligibility. *American Economic Journal: Economic Policy*, 3(2):41–65.

Caylor, K. K., Manfreda, S., and Rodriguez-Iturbe, I. (2005). On the coupled geomorphological and ecohydrological organization of river basins. *Advances in Water Resources*, 28(1):69–86.

Challinor, A. J., Ewert, F., Arnold, S., Simelton, E., and Fraser, E. (2009). Crops and climate change: Progress, trends, and challenges in simulating impacts and informing adaptation. *Journal of experimental botany*, 60(10):2775–2789.

Cheney, C. (2018). Want to reach the world's poorest? Design for dumb phones. `https://www.devex.com/news/want-to-reach-the-world-s-poorest-design-for-dumb-phones-90993`. Accessed: 2019-11-05.

Choi, H. S., Schneider, U. A., Rasche, L., Cui, J., Schmid, E., and Held, H. (2015). Potential effects of perfect seasonal climate forecasting on agricultural markets, welfare and land use: A case study of Spain. *Agric. Syst.*, 133:177–189.

Clapp, R. B. and Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water resources research*, 14(4):601–604.

Communications Authority of Kenya (2018). Third quarterly sector statistics report for the financial year 2017/2018 (1st January - 31st March 2018). Technical report, Communications Authority of Kenya, Nairobi, Kenya.

Crandall, A. (2012). Kenyan farmers' use of cell phones: Calling preferred over SMS. *Proceedings of M4D 2012 28-29 February 2012 New Delhi, India*, 28(29):119.

Cressman, K. (2013). Role of remote sensing in desert locust early warning. *Journal of Applied Remote Sensing*, 7(1):075098.

CSO (2004). Living conditions monitoring survey report 2002–2003. *Survey report, Central Statistical Office*.

Davenport, F. M., Harrison, L., Shukla, S., Husak, G., Funk, C., and McNally, A. (2019). Using out-of-sample yield forecast experiments to evaluate which earth observation products best indicate end of season maize yields. *Environmental Research Letters*, 14(12):124095.

Dinar, A., Hassan, R., Mendelsohn, R., and Benhin, J. (2008). *Climate change and agriculture in Africa: impact assessment and adaptation strategies*. Routledge.

Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A., and Maher, N. (2016). More extreme precipitation in the world's dry and wet regions. *Nature Climate Change*, 6(5):508–513.

Donovan, K. (2017). Anytime, anywhere: Mobile devices and services and their impact on agriculture and rural development. In *ICT in Agriculture (Updated Edition): Connecting Smallholders to Knowledge, Networks, and Institutions*. World Bank.

Downing, T. E., Ringius, L., Hulme, M., and Waughray, D. (1997). Adapting to Climate Change in Africa. *Mitigation and Adaptation Strategies for Global Change*, 2(1):19–44.

Eakin et al. (2015). Information and communication technologies and climate change adaptation in Latin America and the Caribbean: A framework for action. *Climate and Development*, 7(3):208–222.

Erenstein, O., Kassie, G. T., Langyintuo, A. S., and Mwangi, W. (2011). *Characterization of maize producing households in drought prone regions of Eastern Africa*. CIMMYT.

Fafchamps, M. and Minten, B. (2012). Impact of SMS-based agricultural information on Indian farmers. *The World Bank Economic Review*, 26(3):383–414.

Fawcett, D., Pearce, T., Ford, J. D., and Archer, L. (2017). Operationalizing longitudinal approaches to climate change vulnerability assessment. *Global environmental change*, 45:79–88.

Feder, G., Just, R. E., and Zilberman, D. (1985). Adoption of agricultural innovations in developing countries: A survey. *Economic development and cultural change*, 33(2):255–298.

Feder, G. and Umali, D. L. (1993). The adoption of agricultural innovations: A review. *Technological forecasting and social change*, 43(3-4):215–239.

FEWS NET (2014). Zambia livelihood zones and descriptions. Technical report, Famine Early Warning Systems Network, Washington DC.

Fida, M.-R. and Marina, M. K. (2019). Uncovering mobile infrastructure in developing countries with crowdsourced measurements. In *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development*, pages 1–11.

Fischer, E. and Qaim, M. (2012). Linking smallholders to markets: Determinants and impacts of farmer collective action in Kenya. *World Development*, 40(6):1255–1268.

Franz, T. E., Caylor, K. K., Nordbotten, J. M., Rodríguez-Iturbe, I., and Celia, M. A. (2010). An ecohydrological approach to predicting regional woody species distribution patterns in dryland ecosystems. *Advances in Water Resources*, 33(2):215–230.

Funk, C., Davenport, F., Eilerts, G., Nourey, N., and Galu, G. (2018a). Contrasting Kenyan resilience to food insecurity: 2011 and 2017. Technical report, Famine Early Warning Systems Network Report, USAID, Washington D.C.

Funk, C., Harrison, L., Shukla, S., Pomposi, C., Galu, G., Korecha, D., Husak, G., Magadzire, T., Davenport, F., Hillbruner, C., et al. (2018b). Examining the role of unusually warm Indo-Pacific sea-surface temperatures in recent African droughts. *Quarterly Journal of the Royal Meteorological Society*, 144:360–383.

Funk, C., Shukla, S., Thiaw, W. M., Rowland, J., Hoell, A., McNally, A., Husak, G., Novella, N., Budde, M., Peters-Lidard, C., et al. (2019). Recognizing the famine early warning systems network: Over 30 years of drought early warning science advances and partnerships promoting global food security. *Bulletin of the American Meteorological Society*, 100(6):1011–1027.

Funk, C. C. and Brown, M. E. (2009). Declining global per capita agricultural production and warming oceans threaten food security. *Food Security*, 1(3):271–289.

Gates Foundation (2020). Smart farming innovations for small-scale producers: A grand challenge request for proposal. Technical report, Bill and Melinda Gates Foundation, Seattle, Washington.

Gefen, D. and Straub, D. W. (1997). Gender differences in the perception and use of e-mail: An extension to the technology acceptance model. *MIS quarterly*, pages 389–400.

Government of Kenya (2010). *Agricultural Sector Development Strategy, 2010-2020*. Government of Kenya.

Gower, D. B., Dell'Angelo, J., McCord, P. F., Caylor, K. K., and Evans, T. P. (2016). Modeling ecohydrological dynamics of smallholder strategies for food production in dryland agricultural systems. *Environmental Research Letters*, 11(11):115005.

GSMA (2016). The mobile economy. Africa 2016. Technical report, GSMA, London, UK.

Guan, K., Sultan, B., Biasutti, M., Baron, C., and Lobell, D. B. (2017). Assessing climate adaptation options and uncertainties for cereal systems in West Africa. *Agricultural and Forest Meteorology*, 232:291–305.

Guido, Z., Zimmer, A., Lopus, S., Hannah, C., Gower, D., Waldman, K., Krell, N., Sheffield, J., Caylor, K., and Evans, T. (2020). Farmer forecasts: Impacts of seasonal rainfall expectations on agricultural decision-making in Sub-Saharan Africa. *Climate Risk Management*, 30:100247.

Hansen, J., Hellin, J., Rosenstock, T., Fisher, E., Cairns, J., Stirling, C., Lamanna, C., van Etten, J., Rose, A., and Campbell, B. (2019). Climate risk management and rural poverty reduction. *Agricultural Systems*, 172:28–46.

Hansen, J. W., Baethgen, W. E., Osgood, D. E., Ceccato, P. N., and Ngugi, R. K. (2007). Innovations in climate risk management: Protecting and building rural livelihoods in a variable and changing climate. *Journal of Semi-Arid Tropical Agricultural Research*, 4.

Hansen, J. W., Mason, S. J., Sun, L., and Tall, A. (2011). Review of seasonal climate forecasting for agriculture in Sub-Saharan Africa. *Experimental Agriculture*, 47(2):205–240.

Harrison, L., Funk, C., and Peterson, P. (2019). Identifying changing precipitation extremes in Sub-Saharan Africa with gauge and satellite products. *Environmental Research Letters*, 14(8):085007.

Herrero, M., Ringler, C., van de Steeg, J., Thornton, P. K., Zhu, T., Bryan, E., Omolo, A., Koo, J., and Notenbaert, A. M. O. (2010). Climate variability and climate change and their impacts on Kenya's agricultural sector. *ILRI*.

Highet, C., Skelly, H., and Tyers, A. (2017). Gender and information technology (ICT) survey toolkit. Technical report, USAID, Washington, D.C.

IPCC (2007). *Climate change 2007: Impacts, adaptation and vulnerability: contribution of working group II to the fourth assessment report of the IPCC*, volume 4. Cambridge University Press.

IPCC (2014). *Climate change 2014: Synthesis report. Summary for policymakers*. IPCC.

Jahnke, H. E. (1982). *Livestock production systems and livestock development in tropical Africa*, volume 35. Kieler Wissenschaftsverlag Vauk.

Jayne, T. S., Chamberlin, J., Traub, L., Sitko, N., Muyanga, M., Yeboah, F. K., Anseeuw, W., Chapoto, A., Wineman, A., Nkonde, C., et al. (2016). Africa's changing farm size distribution patterns: The rise of medium-scale farms. *Agricultural Economics*, 47(S1):197–214.

Kabechani, A., Deka, B., and Mwila, B. S. (2017). Energy policy reform: The impact of removal of electricity subsidies on small, medium sized and poor households. Technical report, Policy Brief by Policy Monitoring and Research Centre, Lusaka, Zambia.

Kalanda-Joshua, M., Ngongondo, C., Chipeta, L., and Mpembeka, F. (2011). Integrating indigenous knowledge with conventional science: Enhancing localised climate and weather forecasts in Nessa, Mulanje, Malawi.

Karlsson et al. (2017). Accelerating affordable smartphone ownership in emerging markets. Technical report, GSMA, London, UK.

Kassie, M., Jaleta, M., Shiferaw, B., Mmbando, F., and Mekuria, M. (2013). Adoption of interrelated sustainable agricultural practices in smallholder systems: Evidence from rural Tanzania. *Technological forecasting and social change*, 80(3):525–540.

Katul, G., Porporato, A., and Oren, R. (2007). Stochastic dynamics of Plant-Water interactions. *Annu. Rev. Ecol. Evol. Syst.*

Katz, E. and Barandun, A. (2002). *Innovative approaches to financing extension for agriculture and natural resource management*. LBL, Swiss Center for Agricultural Extension, LBL, Swiss Center for Agricultural Extension.

Kebede, Y. (1992). Risk taking behaviour and new technologies: The case of producers in the central highlands of Ethiopia. *Quarterly Journal of International Agriculture*, 31:269–269.

Kenya Bureau of Statistics Economic Survey (2017). Economic Survey 2017. Technical report, National Bureau of Statistics, Nairobi, Kenya.

Kilavi, M., MacLeod, D., Ambani, M., Robbins, J., Dankers, R., Graham, R., Titley, H., Salih, A. A., and Todd, M. C. (2018). Extreme rainfall and flooding over central kenya including Nairobi city during the long-rains season 2018: Causes, predictability, and potential for early warning and actions. *Atmosphere*, 9(12):472.

Kipkemoi, I., Michaelides, K., Rosolem, R., and Singer, M. B. (2021). Climatic expression of rainfall on soil moisture dynamics in drylands. *Hydrology and Earth System Sciences Discussions*, pages 1–24.

Kotir, J. H. (2011). Climate change and variability in Sub-Saharan Africa: A review of current and future trends and impacts on agriculture and food security. *Environment, Development and Sustainability*, 13(3):587–605.

Krell, N., Giroux, S., Guido, Z., Hannah, C., Lopus, S., Caylor, K., and Evans, T. (2020). Smallholder farmers' use of mobile phone services in central Kenya. *Climate and Development*, pages 1–13.

Kurukulasuriya, P., Mendelsohn, R., Hassan, R., Benhin, J., Deressa, T., Diop, M., Eid, H. M., Fosu, K. Y., Gbetibouo, G., Jain, S., et al. (2006). Will African agriculture survive climate change? *The World Bank Economic Review*, 20(3):367–388.

Laio, F., Porporato, A., Fernandez-Illescas, C. P., and Rodríguez-Iturbe, I. (2001a). Plants in water-controlled ecosystems: Active role in hydrologic processes and response to water stress: IV. Discussion of real cases. *Adv. Water Resour.*, 24(7):745–762.

Laio, F., Porporato, A., Ridolfi, L., and Rodríguez-Iturbe, I. (2001b). Plants in water-controlled ecosystems: Active role in hydrologic processes and response to water stress: II. Probabilistic soil moisture dynamics. *Advances in water resources*, 24(7):707–723.

Leenaars, J. (2014). Africa Soil Profiles Database, Version 1.2. A compilation of geo-referenced and standardized legacy soil profile data for Sub Saharan Africa (with dataset). `https://data.isric.org/geonetwork/srv/eng/catalog.search#`

`/metadata/2a7d2fb8-e0db-4a4b-9661-4809865aaccf`. Wageningen, the Netherlands: Africa Soil Information Service (AfSIS) project and ISRIC—World Soil Information.

Liebmann, B., Hoerling, M. P., Funk, C., Bladé, I., Dole, R. M., Allured, D., Quan, X., Pegion, P., and Eischeid, J. K. (2014). Understanding recent eastern Horn of Africa rainfall variability and change. *Journal of Climate*, 27(23):8630–8645.

Liniger, H. and Thomas, D. B. (1998). Grass: Ground cover for restoration of arid and semi-arid soils. *Advances in GeoEcology*, 31:1167–1178.

Liniger, H. and Weingartner, R. (1998). Mountains and freshwater supply. *Unasylva (FAO)*.

Lobell, D. B., Schlenker, W., and Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. *Science*, 333(6042):616–620.

Lopus, S., McCord, P., Gower, D., and Evans, T. (2017). Drivers of farmer satisfaction with small-scale irrigation systems. *Applied Geography*, 89:77–86.

Lyon, B. and DeWitt, D. G. (2012). A recent and abrupt decline in the East African long rains. *Geophysical Research Letters*, 39(2).

Marenya, P. P. and Barrett, C. B. (2007). Household-level determinants of adoption of improved natural resources management practices among smallholder farmers in western Kenya. *Food policy*, 32(4):515–536.

Markelova, H. and Mwangi, E. (2010). Collective action for smallholder market access: Evidence and implications for Africa. *Review of policy research*, 27(5):621–640.

Mbow, C., Rosenzweig, C., Barioni, L., Benton, T., Herrero, M., Krishnapillai, M., Liwenga, E., Pradhan, P., Rivera-Ferre, M., Sapkota, T., et al. (2019). Food security, chapter 5. *Special Report on Climate Change and Land. IPCC: Intergovernmental Panel on Climate Change.*

McCarthy, N., Lipper, L., Branca, G., et al. (2011). Climate-smart agriculture: Smallholder adoption and implications for climate change adaptation and mitigation. *Mitigation of Climate Change in Agriculture Working Paper*, 3:1–37.

McCord, P., Dell'angelo, J., Gower, D., Caylor, K., and Evans, T. (2017). Household-level heterogeneity of water resources within common-pool resource systems. *Ecology and Society*, 22(1).

McCord, P., Waldman, K., Baldwin, E., Dell'Angelo, J., and Evans, T. (2018). Assessing multi-level drivers of adaptation to climate variability and water insecurity in smallholder irrigation systems. *World Development*, 108:296–308.

McNutt, P. (1999). Public goods and club goods. *Encyclopedia of law and economics*, 1:927–951.

Meehl, G. A., Stocker, T. F., Collins, W. D., Friedlingstein, P., Gaye, A. T., Gregory, J. M., Kitoh, A., Knutti, R., Murphy, J. M., Noda, A., et al. (2007). Global climate projections. Climate Change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.

Meso, P., Musa, P., and Mbarika, V. (2005). Towards a model of consumer use of mobile information and communication technology in LDCs: The case of sub-Saharan Africa. *Information Systems Journal*, 15(2):119–146.

Mittal, S., Gandhi, S., and Tripathi, G. (2010). Socio-economic impact of mobile phones on Indian agriculture. Technical report, Working paper, No. 246, Indian Council for Research on International Economic Relations (ICRIER), New Delhi.

Mittal, S. and Mehar, M. (2016). Socio-economic factors affecting adoption of modern information and communication technology by farmers in India: Analysis using multivariate probit model. *The Journal of Agricultural Education and Extension*, 22(2):199–212.

Morris, M. (2001). Assessing the benefits of international maize breeding research: An overview of the global maize impacts study. Part II of the CIMMYT 1999–2000 world maize facts and trends. Technical report, CIMMYT, Mexico City.

Morton, J. F. (2007). The impact of climate change on smallholder and subsistence agriculture. *Proceedings of the national academy of sciences*, 104(50):19680–19685.

Moser, C. O. (1998). The asset vulnerability framework: Reassessing urban poverty reduction strategies. *World development*, 26(1):1–19.

Muchena, F. and Gachene, C. (1988). Soils of the highland and mountainous areas of Kenya with special emphasis on agricultural soils. *Mountain Research and Development*, pages 183–191.

Mudenda, C., Johnson, D., Parks, L., and van Stam, G. (2013). Power instability in rural Zambia, case Macha. In *International Conference on e-Infrastructure and e-Services for Developing Countries*, pages 260–270. Springer.

Muller, C., Cramer, W., Hare, W. L., and Lotze-Campen, H. (2011). Climate change risks for African agriculture. *Proceedings of the National Academy of Sciences*, 108(11):4313–4315.

Muyanga, M. and Jayne, T. (2008). Private agricultural extension system in Kenya: Practice and policy lessons. *Journal of agricultural education and extension*, 14(2):111–124.

Mwesige, P. G. (2004). Cyber elites: A survey of internet café users in Uganda. *Telematics and Informatics*, 21(1):83–101.

Naidoo, S., London, L., Burdorf, A., Naidoo, R. N., and Kromhout, H. (2008). Agricultural activities, pesticide use and occupational hazards among women working in small scale farming in Northern KwaZulu-Natal, South Africa. *International journal of occupational and environmental health*, 14(3):218–224.

Nakalembe, C. (2018). Characterizing agricultural drought in the Karamoja subregion of Uganda with meteorological and satellite-based indices. *Natural Hazards*, 91(3):837–862.

Nakalembe, C. (2020). Urgent and critical need for sub-saharan African countries to invest in earth observation-based agricultural early warning and monitoring systems. *Environmental Research Letters*, 15(12):121002.

Nakasone, E., Torero, M., and Minten, B. (2014). The Power of Information: The ICT Revolution in Agricultural Development. *Annual Review of Resource Economics*, 6(1):533–550.

Narrod et al. (2009). Public–private partnerships and collective action in high value fruit and vegetable supply chains. *Food policy*, 34(1):8–15.

Nations, U. (2017). World population prospects: The 2017 revision, volume I: Comprehensive tables (ST/ESA/SER.A/399). Technical report, United Nations, Department of Economic and Social Affairs, Population Division, New York, USA.

Neumann, P. M. (2008). Coping mechanisms for crop plants in drought-prone environments. *Annals of Botany*, 101(7):901–907.

Nicholson, S. E. (2017). Climate and climatic variability of rainfall over eastern Africa. *Reviews of Geophysics*, 55(3):590–635.

Niles, M. T. and Brown, M. E. (2017). A multi-country assessment of factors related to smallholder food security in varying rainfall conditions. *Scientific Reports*, 7(1):16277.

Nyakudya, I. W. and Stroosnijder, L. (2014). Effect of rooting depth, plant density and planting date on maize (Zea mays L.) yield and water use efficiency in semi-arid Zimbabwe: Modelling with AquaCrop. *Agricultural Water Management*, 146:280–296.

Ogalleh, S., Vogl, C., Eitzinger, J., and Hauser, M. (2012). Local perceptions and responses to climate change and variability: The case of Laikipia district, Kenya. *Sustainability*, 4(12):3302–3325.

Ogutu, S. O., Okello, J. J., and Otieno, D. J. (2014). Impact of information and communication technology-based market information services on smallholder farm input use and productivity: The case of Kenya. *World Development*, 64:311–321.

OpenCellID (2018a). Data sources. `http://wiki.opencellid.org/wiki/Data_sources`, Last accessed on 2021-04-29.

OpenCellID (2018b). Frequently asked questions. `http://wiki.opencellid.org/wiki/FAQ#Calculated_cell_position_vs_exact_cell_position`, Last accessed on 2021-04-29.

Patakamuri, S. K. and O'Brien, N. (2020). *modifiedmk: Modified Versions of Mann Kendall and Spearman's Rho Trend Tests*. R package version 1.5.0.

Patt, A., Suarez, P., and Gwata, C. (2005). Effects of seasonal climate forecasts and participatory workshops among subsistence farmers in Zimbabwe. *Proceedings of the National Academy of Sciences*, 102(35):12623–12628.

Piccoli, G., Ahmad, R., and Ives, B. (2001). Web-based virtual learning environments: A research framework and a preliminary assessment of effectiveness in basic IT skills training. *MIS quarterly*, pages 401–426.

Porporato, A., D'Odorico, P., Laio, F., Ridolfi, L., and Rodríguez-Iturbe, I. (2002). Ecohydrology of water-controlled ecosystems. *Advances in Water Resources*, 25(8-12):1335–1348.

Porporato, A., Laio, F., Ridolfi, L., Caylor, K. K., and Rodriguez-Iturbe, I. (2003). Soil moisture and plant stress dynamics along the Kalahari precipitation gradient. *Journal of Geophysical Research: Atmospheres*, 108(D3).

Porporato, A., Laio, F., Ridolfi, L., and Rodríguez-Iturbe, I. (2001). Plants in water-controlled ecosystems: Active role in hydrologic processes and response to water stress: III. vegetation water stress. *Adv. Water Resour.*, 24(7):725–744.

Porter, G., Hampshire, K., Abane, A., Munthali, A., Robson, E., Mashiri, M., and Tanle, A. (2012). Youth, mobility and mobile phones in Africa: Findings from a three-country study. *Information Technology for Development*, 18(2):145–162.

Python Software Foundation (2019). Python language reference.

Qiang, C. Z., Kuek, S. C., Dymond, A., and Esselaar, S. (2012). Mobile applications for agriculture and rural development. Technical report, World Bank, Washington, DC.

Qualtrics (2019). Qualtrics.

Ray, D. K., Gerber, J. S., Macdonald, G. K., and West, P. C. (2015). Climate variation explains a third of global crop yield variability. *Nature Communications*, 6:1–9.

Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C., and Foley, J. A. (2012). Recent patterns of crop yield growth and stagnation. *Nature communications*, 3:1293.

Ray, D. K., West, P. C., Clark, M., Gerber, J. S., Prishchepov, A. V., and Chatterjee, S. (2019). Climate change has likely already affected global food production. *PloS one*, 14(5):e0217148.

Recha, C., Makokha, G., Traore, P., Shisanya, C., Lodoun, T., and Sako, A. (2012). Determination of seasonal rainfall variability, onset and cessation in semi-arid Tharaka district, Kenya. *Theoretical and Applied Climatology*, 108(3-4):479–494.

Reidsma, P., Ewert, F., Lansink, A. O., and Leemans, R. (2010). Adaptation to climate change and climate variability in european agriculture: the importance of farm level responses. *European journal of agronomy*, 32(1):91–102.

Ricciardi, V., Mehrabi, Z., Wittman, H., James, D., and Ramankutty, N. (2021). Higher yields and more biodiversity on smaller farms. *Nature Sustainability*, pages 1–7.

Rigden, A., Mueller, N., Holbrook, N., Pillai, N., and Huybers, P. (2020). Combined influence of soil moisture and atmospheric evaporative demand is important for accurately predicting US maize yields. *Nature Food*, 1(2):127–133.

Roche, K. R., Müller-Itten, M., Dralle, D. N., Bolster, D., and Müller, M. F. (2020). Climate change and the opportunity cost of conflict. *Proceedings of the National Academy of Sciences*, 117(4):1935–1940.

Rodríguez-Iturbe, I. and Porporato, A. (2007). *Ecohydrology of water-controlled ecosystems: Soil moisture and plant dynamics*. Cambridge University Press.

Rodríguez-Iturbe, I., Porporato, A., Laio, F., and Ridolfi, L. (2001). Plants in water-controlled ecosystems: Active role in hydrologic processes and response to water stress: I. Scope and general outline. *Adv. Water Resour.*, 24(7):695–705.

Rotondi, V., Kashyap, R., Pesando, L. M., Spinelli, S., and Billari, F. C. (2020). Leveraging mobile phones to attain sustainable development. *Proceedings of the National Academy of Sciences*, 117(24):13413–13420.

Roudier, P., Muller, B., d'Aquino, P., Roncoli, C., Soumaré, M. A., Batté, L., and Sultan, B. (2014). The role of climate forecasts in smallholder agriculture: Lessons from participatory research in two communities in Senegal. *Climate Risk Management*, 2:42–55.

Sah, R., Chakraborty, M., Prasad, K., Pandit, M., Tudu, V., Chakravarty, M., Narayan, S., Rana, M., and Moharana, D. (2020). Impact of water deficit stress in maize: Phenology and yield components. *Scientific reports*, 10(1):1–15.

Salgado-Aguilar, M., Molnar, T., Pons-Hernández, J. L., Covarrubias-Prieto, J., Ramírez-Pimentel, J. G., Raya-Pérez, J. C., Hearne, S., and Iturriaga, G. (2020). Physiological and biochemical analyses of novel drought-tolerant maize lines reveal osmoprotectant accumulation at silking stage. *Chilean journal of agricultural research*, 80(2):241–252.

Samberg, L. H., Gerber, J. S., Ramankutty, N., Herrero, M., and West, P. C. (2016). Subnational distribution of average farm size and smallholder contributions to global food production. *Environmental Research Letters*, 11(12):124010.

Santosham, S. and Lindsey, D. (2015). Bridging the gender gap: Mobile access and usage in low-and middle-income countries. Technical report, GSMA, London, UK.

Schmocker, J., Liniger, H., Ngeru, J., Brugnara, Y., Auchmann, R., and Brönnimann, S. (2016). Trends in mean and extreme precipitation in the mount Kenya region from observations and reanalyses. *International journal of climatology*, 36(3):1500–1514.

Senay, G. (2004). Crop Water Requirement Satisfaction Index (WRSI) Model Description.

Setimela, P. S., Magorokosho, C., Lunduka, R., Gasura, E., Makumbi, D., Tarekegne, A., Cairns, J. E., Ndhlela, T., Erenstein, O., and Mwangi, W. (2017). On-farm yield gains with stress-tolerant maize in eastern and southern Africa. *Agronomy Journal*, 109(2):406–417.

Shongwe, M. E., van Oldenborgh, G. J., van den Hurk, B., and van Aalst, M. (2011). Projected changes in mean and extreme precipitation in Africa under global warming. Part II: East Africa. *Journal of climate*, 24(14):3718–3733.

Singer, M. B. and Michaelides, K. (2017). Deciphering the expression of climate change within the Lower Colorado River basin by stochastic simulation of convective rainfall. *Environmental Research Letters*, 12(10):104011.

Slingo, J. M., Challinor, A. J., Hoskins, B. J., and Wheeler, T. R. (2005). Introduction: Food crops in a changing climate. *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Smale, M. and Jayne, T. S. (2010). ‘Seeds of success’ in retrospect: Hybrid maize in eastern and southern Africa. *Successes in African agriculture: Lessons for the future*, pages 71–112.

Smit, B. and Wandel, J. (2006). Adaptation, adaptive capacity and vulnerability. *Global environmental change*, 16(3):282–292.

Solomon, S., Manning, M., Marquis, M., Qin, D., et al. (2007). *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC*, volume 4. Cambridge university press.

Song, S. (2017). Mapping the underserved. `https://manypossibilities.net/2017/04/mapping-the-unserved/`, Last accessed on 2021-04-27.

Souter et al. (2005). The economic impact of telecommunications on rural livelihoods and poverty reduction: A study of rural communities in india (gujarat), mozambique and tanzania. *IIMA Working Papers WP2005-11-04*.

Srinivasan, J. and Burrell, J. (2013). Revisiting the fishers of Kerala, India. In *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers-Volume 1*, pages 56–66. ACM.

Sweeney, S. and Arabadjis, S. (2021). *Spatial Point Patterns in the Social Sciences*. Chapter submitted for review.

Sweeney, S. and Gómez-Antonio, M. (2016). Localization and industry clustering econometrics: An assessment of gibbs models for spatial point processes. *Journal of Regional Science*, 56(2):257–287.

Taylor, R. G., Todd, M. C., Kongola, L., Maurice, L., Nahozya, E., Sanga, H., and MacDonald, A. M. (2013). Evidence of the dependence of groundwater resources on extreme rainfall in east Africa. *Nature Climate Change*, 3(4):374–378.

Tesfaye, K., Sonder, K., Caims, J., Magorokosho, C., Tarekegn, A., Kassie, G. T., Getaneh, F., Abdoulaye, T., Abate, T., and Erenstein, O. (2016). Targeting drought-tolerant maize varieties in southern Africa: A geospatial crop modeling approach using big data. *International Food and Agribusiness Management Review IFAMA*, 19:75–92.

Thiga, M. and Ndungu, S. (2015). Utilization of ICT for agriculture: A case study of Kakamega county, Kenya. In *2015 IST-Africa Conference*, pages 1–9. IEEE.

Thornton, P. K., Jones, P. G., Alagarswamy, G., Andresen, J., and Herrero, M. (2010). Adapting to climate change: Agricultural system and household impacts in east Africa. *Agricultural systems*, 103(2):73–82.

Thornton, P. K., Jones, P. G., Ericksen, P. J., and Challinor, A. J. (2011). Agriculture and food systems in sub-Saharan Africa in a 4C+ world. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 369(1934):117–36.

Trenberth, K. E. (2011). Changes in precipitation with climate change. *Climate Research*, 47(1-2):123–138.

Ulm, M., Widhalm, P., and Brändle, N. (2015). Characterization of mobile phone localization errors with OpenCellid data. In *2015 4th International Conference on Advanced Logistics and Transport (ICALT)*, pages 100–104. IEEE.

UN Women (2015). Progress of the world's women 2015-2016: Transforming economies, realizing rights. Technical report, United Nations Entity for Gender Equality and the Empowerment of Women.

USAID (2020). Zambia power Africa fact sheet. `https://www.usaid.gov/powerafrica/zambia`, Last accessed on 2021-04-21.

Van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., and Hochman, Z. (2013). Yield gap analysis with local to global relevance—a review. *Field Crops Research*, 143:4–17.

Vergopolan, N., Xiong, S., Estes, L., Wanders, N., Chaney, N. W., Wood, E. F., Konar, M., Caylor, K., Beck, H. E., Gatti, N., et al. (2021). Field-scale soil moisture bridges the spatial-scale gap between drought monitoring and agricultural yields. *Hydrology and Earth System Sciences*, 25(4):1827–1847.

Vervoort, R. W., Willem Vervoort, R., Muita, R., Ampt, P., and van Ogtrop, F. (2016). Managing the water cycle in Kenyan small-scale maize farming systems: Part 2. Farmers' use of formal and informal climate forecasts. *Wiley Interdisciplinary Reviews: Water*, 3(1):127–140.

Waldman, K. B., Vergopolan, N., Attari, S. Z., Sheffield, J., Estes, L. D., Caylor, K. K., and Evans, T. P. (2019). Cognitive biases about climate variability in smallholder farming systems in Zambia. *Weather, Climate, and Society*, 11(2):369–383.

Waters, J. and Adger, W. N. (2017). Spatial, network and temporal dimensions of the determinants of adaptive capacity in poor urban areas. *Global environmental change*, 46:42–49.

WFP and FAO (2021). Hunger hotspots: FAO-WFP early warnings on acute food insecurity: March to July 2021 outlook. Technical report, Food and agricultural organization of the United Nations (FAO), Rome, Italy.

Wiesmann, U. (1998). *Sustainable regional development in rural Africa: Conceptual framework and case studies from Kenya*. University of Berne Switzerland, Institute of Geography.

Williams, A. P. and Funk, C. (2011). A westward extension of the warm pool leads to a westward extension of the Walker circulation, drying eastern Africa. *Climate Dynamics*, 37(11-12):2417–2435.

Williams, C. A. and Albertson, J. D. (2004). Soil moisture controls on canopy-scale water and carbon fluxes in an African savanna. *Water Resources Research*, 40(9).

Wood, S. A., Jina, A. S., Jain, M., Kristjanson, P., and DeFries, R. S. (2014). Smallholder farmer cropping decisions related to climate variability across multiple regions. *Glob. Environ. Change*, 25:163–172.

World Bank (2017). Environmental and social management framework Zambia electricity service access project. no p162760, report number sfg3353. Technical report, World Bank, Washington DC.

World Bank (2018a). Information and Communications for Development 2018: Data-Driven Development. Technical report, World Bank, Washington, DC.

World Bank (2018b). Tracking SDG7: The energy progress report 2018. Technical report, International Energy Agency; International Renewable Energy Agency; United Nations; World Bank Group; World Health Organization, World Bank, Washington DC.

Wyche, S. and Murphy, L. (2012). Dead China-make phones off the grid: investigating and designing for mobile phone use in rural Africa. In *Proceedings of the Designing Interactive Systems Conference*, pages 186–195. ACM.

Wyche, S. and Olson, J. (2018). Gender, mobile, and mobile internet— Kenyan women's rural realities, mobile internet access, and "Africa rising". *Information Technologies & International Development*, 14:15.

Wyche, S., Simiyu, N., and Othieno, M. E. (2019). Understanding women's mobile phone use in rural Kenya: An affordance-based approach. *Mobile Media & Communication*, 7(1):94–110.

Wyche, S. and Steinfield, C. (2016). Why don't farmers use cell phones to access market prices? Technology affordances and barriers to market information services adoption in rural Kenya. *Information Technology for Development*, 22(2):320–333.

Yue, S. and Wang, C. (2004). The mann-kendall test modified by effective sample size to detect trend in serially correlated hydrological series. *Water resources management*, 18(3):201–218.

Zheng, J., Fan, J., Zhang, F., Yan, S., and Xiang, Y. (2018). Rainfall partitioning into throughfall, stemflow and interception loss by maize canopy on the semi-arid Loess Plateau of China. *Agricultural water management*, 195:25–36.

Ziervogel, G., Bithell, M., Washington, R., and Downing, T. (2005). Agent-based social simulation: A method for assessing the impact of seasonal climate forecast applications among smallholder farmers. *Agric. Syst.*, 83(1):1–26.