ASSESSING AND MODELING CROP YIELD AND SOIL CARBON IN SMALLHOLDER FIELDS IN AFRICA AND CENTRAL AMERICA

By

Lin Liu

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Environmental Geosciences—Doctor of Philosophy
Environmental Science and Policy – Dual Major

2020

ABSTRACT

ASSESSING AND MODELING CROP YIELD AND SOIL CARBON IN SMALLHOLDER FIELDS IN AFRICA AND CENTRAL AMERICA

By

Lin Liu

Research on developing and testing agricultural tools for smallholder agricultural producers remains limited , despite the fact that new tools designed to improve agronomic management and resilience of cropping systems are becoming increasingly available also in developing countries. The overarching goal of this dissertation was to evaluate the efficacy of agricultural technologies, like process-based crop simulation models, to improve food production forecasts and crop yields in smallholder fields in low-income countries (Chapter 1).

Chapter 2 presents the development and the validation of a new maize yield forecasting system for the Government of Tanzania. In this study, a field-based survey was integrated with a process-based crop model, Systems Approach to Land Use Sustainability (SALUS) to provide accurate and timely maize yield forecasts for small fields in Tanzania. In spite of a wide range of maize growing conditions, the method developed in the chapter has shown to provide reliable forecasts across three districts in Tanzania 14-77 days prior to crop harvest.

Chapter 3 investigates how climate impact assessment differs when using the averaged value simulated with each climate model from the Coordinated Regional Climate Downscaling Experiment (CORDEX) program, versus one simulated value with one single delta-method projected climate based on average changes in climatic variables. This analysis was performed using SALUS-simulated grain yield, Soil Organic Caron (SOC) and soil inorganic Nitrogen (N) for 60 sites from Chapter 2. The simulated climate impact on soil N and SOC using the delta-method climate was close to the average simulated impact using each climate model, but the

adverse impact on grain yield was projected to be lower.

Chapter 4 focuses on agronomic management that could increase the yield of yam while improving soil fertility in Ghana. In this study, I first parameterized yam in the SALUS model using field experiment data from three N and phosphorous (P) fertilizer treatments combined with two yam cropping in two distinct agroecological zones in Ghana for two years. The calibrated and validated SALUS-Yam model was used to assess the impact of four management treatments: continuous unfertilized rainfed yam (control), pigeonpea-yam rotation, yam with 3 Mg/ha pigeonpea residue incorporated, and yam with 23-23 kg/ha N-P_2O_5 fertilizer added. The results showed that incorporating pigeonpea residues into yam fields produced the highest yam tuber yield and reduced SOC compared to the other treatments. This work also confirmed that yam cultivation in Ghana was mostly limited by lack of nutrients (N, P or both), as opposed to drought.

Chapter 5 presents within-field variability of smallholder fields. From field observations in Tanzania, maize-based fields across more than 60 sites in three districts in Tanzania contained considerable variability in plant density (median CV 20-30%) and grain yield (median CV 30-36%). Grain yield variability was correlated with in-season vegetation indices, particularly the green chlorophyll vegetation index. The coefficient of variation of normalized difference vegetation index became smaller as the spatial resolution became coarser. The analysis was performed using images of distinct spatial resolutions for smallholder yam and pigeonpea fields in Ghana, and bean growing areas in Honduras.

Lessons from the research projects and recommendations on using agricultural technologies for international agricultural development are outlined in Chapter 6.

To my parents, my friends, and my advisor Dr. Bruno Basso.

ACKNOWLEDGEMENTS

It takes a village to complete a PhD program. I had an enormous amount of luck to meet with groups of marvelous individuals during the pursuit of my PhD degree at Michigan State University.

First and foremost, I owe who I am now to my PhD advisor and chair of my dissertation committee, Dr. Bruno Basso. Bruno has tapped me with his transformative magic. In the past five years, Bruno has provided a stable space for me to grow. Bruno gives me both freedom and guidance to explore research topics and career potentials. Because of him, I have been exposed to a range of disciplines and tools that can address the grand challenge of feeding the future, and have had the privilege to collaborate with diverse stakeholders to improve food sustainability and security in several countries. I would not be able to complete my degree without his financial investment throughout my time working with him as a graduate student. Bruno's influence on me is far beyond my professional growth. I have learnt important life skills from him, to list a few, working wisely and enjoying life, being positive and adaptive, thinking creatively and asking critical questions, and trusting and forgiving both myself and others. I owe my scholarly and personal achievements to Bruno.

My gratitude goes to my other three dissertation committee members. I am a better scholar because of them. Dr. Jeff Andresen guides me in critically evaluating quality of climate data; his enthusiasm about weather and climate influences me to appreciate all weather events. Dr. David Hyndman provides advice to me on research planning and collaboration, and professional development during my PhD career. He has also been a role model because of his commitment to his two job titles, department chair and principal investigator of his research laboratory. Dr. David Tschirley, who is an expert in international development, ignites my interests in food

systems, and inspires me to examine food security from market and food policy perspectives. The professional growth I gained from interacting with them is invaluable.

This dissertation would not exist without the collaborators in the Government of Tanzania (GoT), United Nations Food and Agriculture organizations (UN FAO), Council for Scientific and Industrial Research-Crops Research Institution (CSIR-CRI) in Ghana, and Michigan State University. Particularly, I would like to thank Wilson Katunzi, Titus Mwisomba and Festo Mwemutsi from the GoT for providing logistical support. I would like to thank Michael Rahija from UN FAO for his technical support in coding, storing, and sharing the field questionnaire surveys. I am grateful for all staff members who were involved in training the enumerators and collecting field data. I thank the enumerators for collecting field data. I would like to thank my colleague and my fellow graduate student Eric Owusu Danquah and his team in CSIR-CRI for collecting and sharing field experiment data about yam development and growth, which was used for model validation in Chapter 4 and image inquiry in Chapter 5.

I thank Bruno Basso's team, including lab technicians, research associates, my graduate peers and visiting scholars, for providing field and programming assistance, and creating a safe environment to work in. My special thanks goes to Dr. Joe Ritchie. I share office space with Dr. Ritchie sometimes when he visits. I am inspired by Dr. Ritchie's lifelong commitment to agronomic research and his dedication to mentoring next-generation researchers.

I received tremendous support from the Department of Earth and Environmental Sciences (EES) and the Environmental Science and Policy Programs (ESPP). My PhD journey was much enriched because of the faculty, the staff members, and the graduate students in the two degree programs. I was able to concentrate on my research because staff members, including Dallas Coryell, Elizabeth McElroy, Ami McMurphy, Pam Robinson, Judi Smelser, and Brittany Walter

from EES, and Marcy Heberer, Sean Lawrie, and Karessa Weir from ESPP, were taking care of travel logistics and degree paperwork. In addition, I would like to express my appreciation for fellowships and travel grants offered by the two degree programs that made it possible for me to present at various international research and development conferences.

I am grateful for my doctors and physicians in East Lansing area, including Drs. G. David Harris, Marcia Kent, Elizabeth and Stephen Phillips, and Julie Matyasic. They helped me to be both physically and mentally fit to undertake challenging, sometimes seemingly impossible, tasks during my PhD program.

I would like to thank all of my friends for their kind support. I always know I am one phone call away from being well when times are rough. I have created warm memories with a group of wonderful people. The list of friends I made during my PhD degree would run long; but I want to acknowledge Jill Deines, Erin Haccker, Dee Jordan, and Yingyue "Juno" Liu – our friendship gives me much strength. In addition, I would like to thank the following individuals who consistently worked with me remotely during the COVID-19 work-from-home era: Hope Akaeze, Tingqiao Chen, Jiaye Guo, Jennifer Lai, and Brittany Tucker.

Last but not the least, my deepest appreciation goes to my mom and my dad. Mom, thank you for your love and caring, and for being a strong woman. Dad, thank you for always prioritizing my education and always trusting my ability in pursuing a career in science. I am grateful that my parents raised me to be an independent, ambitious and diligent woman.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

AfSIS	Africa Soil Information Service
AgMERRA	Agmip Climate Forcing Based On The Modern-Era Retrospective Analysis
AgMIP	Agricultural Model Intercomparison And Improvement Project
APSIM	Agricultural Production Systems sIMulator
CERES	Crop Environment REsource Synthesis
CHIRPS	Climate Hazards Group InfraRed Precipitation With Station
CORDEX	Coordinated Regional Climate Downscaling Experiment
CROPSYST	Cropping Systems Simulation
CV	Coefficient Of Variation
DM	Dry Matter
DSSAT	Decision Support System For Agrotechnology Transfer
EPIC	Environmental Policy Integrated Climate
EVI	Enhanced Vegetation Index
GCVI	Green Chlorophyll Vegetation Index
gNDVI	Green Normalized Difference Vegetation Index
GoT	Government Of Tanzania
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
NASA	National Aeronautics And Space Administration
NDVI	Normalized Difference Vegetation Index
NSE	Nash-Sutcliffe Efficiency
RCP	Representative Concentration Pathway

RMSD	Root Mean Square Of Deviation
SALUS	Systems Approach To Land Use Sustainability
SOC	Soil Organic Carbon
SPAM	Spatial Production Allocation Model
SSA	Sub-Saharan Africa
UAV	Unmanned aerial vehicle
UN FAO	United Nations Food And Agriculture Organizations

CHAPTER 1: INTRODUCTION TO THE DISSERTATION

Rationale and Background

In 2015, the United Nations set 17 Sustainable Development Goals to guide global development agendas. Several of the goals were directly related to the agriculture sector, including ending poverty, achieving zero hunger, producing food sustainably, and combating climate change (Nhemachena et al., 2018). Considering the prevalence of poverty and malnutrition in low- and low-to-middle-income countries, such as countries in Sub-Saharan Africa (SSA) and Honduras, agriculture in these low- and low-to-middle-income countries needs to be reformed to increase land productivity and sustainability (Abraham and Pingali, 2020; Brown et al., 2019).

Agriculture in low- and low-to-middle-income countries is dominated by smallholder farms, where field size is less than 2 ha. These smallholder farms are family operated to supply households with staple food including cereal crops (e.g. maize, rice, and sorghum) and tuber crops (e.g. yam and potatoes) (Lowder et al., 2016). Smallholder fields are often not managed intensely. Food crop cultivation in the small fields is characterized by low planting density, and relies on rain events and nutrient-depleted soils. Crop yield resulting from such environments and management is low (Berre et al., 2017; Tittonell and Giller, 2013). Smallholder farming systems, nonetheless, play a critical role in securing national and household food production.

Food security was first defined as “all people at all times have access to safe and nutritious food to maintain a healthy and active life” in the 1990s by the Food and Agriculture Organization of the United Nations Committee on World Food Security (FAO, 1996). By this definition, food security consists of four building blocks: availability, stability, accessibility, and utilization. Food security is affected by both biophysical and socioeconomic processes. From a

biophysical perspective, having a viable food production system is essential to maintain food availability and stability. For SSA countries, food supply does not solely rely on domestic food production, but is also affected by trade decisions. This entails two prongs of action for protecting food security. At a household level, yields of food crops need to be preserved or improved to provide household income and food supply. At a national level, trade and food policy decisions must be made based on evidence-based agricultural information. One of the important pieces of agricultural information for national food policies is crop production forecasts (Diao, 2016; Jayne and Rashid, 2010). Understanding of future crop yield lays the foundation for both household and national food security.

Growing food crops is not a trivial task. Crop growth and harvestable yield is determined not only by the environment (e.g. rainfall, soil fertility, and temperature) but also by how crops are managed (e.g. irrigation water and fertilizer use). Management practices, perhaps, are the most decisive factor for crop yield in the short term, given that one cannot change weather and can only modify soil conditions through amendments. Soil, climate, and their interactions are critical factors in long-term food production, but they are often neglected in the search for sustainable agricultural systems for feeding the future. Crop yield must be studied from a systems lens that considers the interconnected plant-climate-soil-management continuum. In recent decades, researchers have paid much attention to adverse climate change impact on food availability and stability. Evidence has shown farming systems will be inevitably affected as the climate continues to become warmer and precipitation events become more erratic (Lesk et al., 2016; Serdeczny et al., 2017; Zhao et al., 2017). For SSA countries, though currently the ratio between domestic food production and consumption is 0.8 (80% self-sufficient), crop yield will need a massive increase in order to feed the projected doubling population by 2050, since expanding

agricultural land into current forest land is not a desirable option. Investment in agricultural technologies that can sustainably improve yield is required to achieve the Sustainable Development Goal of zero hunger (Mason-D'Croz et al., 2019; van Ittersum et al., 2016).

At a national level, food policies related to trading decisions rely on timely crop production forecasts. Crop yield forecasts are a key component of the production forecasts. Extensive studies have investigated ways to provide accurate and timely yield forecasts, but most yield-forecasting studies focus on large fields with a single crop per field in the US and European countries. While major trading countries have governmental agencies in charge of issuing national/regional crop yield forecasts throughout a growing season, SSA countries often lack information about seasonal crop status and yield expectations at the end of a season (Gennari and Fonteneau, 2016). Developing an accurate and timely crop yield forecasting system that African countries can implement is an urgent task for achieving the Sustainable Development Goals by 2030.

Agronomic decision-making technologies have been developed to address challenges in crop production. One of the advanced decision-making tools is crop models. Crop models were designed to capture the biophysical interactions between management practices, crop genetics, and growing conditions, including weather and soil. They can be used to test management options to balance yield improvement and environmental conservation and to engage growers about their decision-making (Whitbread et al., 2010). More importantly, crop models can be applied to cropping systems across various climate and soil conditions once they are calibrated and validated for a specific cropping system using field observations (Asseng et al., 2014). Remote sensing technologies, including satellites, micro-satellites, and unmanned aerial vehicles, can provide real-time information about vegetation development and growth over a large spatial

extent. Vegetation indices, derived from reflectance of different wavelength ranges, have been developed and linked to plant health and biomass yield. Advances have been made in integrating crop modeling and remotely sensed images to predict crop yield in SSA countries (Burke and Lobell, 2017; Jin, 2019).

The advanced agricultural technologies, nonetheless, have not been widely adapted or adopted to guide staple food production in Africa. Crop models have been tested globally but only a small fraction of studies have been in African countries (Basso et al., 2016c). Various yield forecasting methodologies have been developed, but only a few investigated the possibility of timely in-season yield forecasts for smallholder fields (Basso and Liu, 2019). Applying remotely sensed images to crop monitoring and management in smallholder fields remains challenging due to field heterogeneity, field size, saturation effect, and cloud cover (Jain 2020; Lambert et al., 2018).

Data availability is another barrier to testing and adapting these advanced agricultural technologies for sustainable crop management in Africa. Crop models need calibration and validation for a given study site before they can be applied to inform management strategies. Remote sensing images need to be verified with ground-truthed data. The use of advanced technology to guide management and improve crop yield entails demand for quality agro-climatic data. This quality agronomic and climatic data, however, does not exist or is not accessible to the public.

We need creative thinking and collaborative efforts to address issues related to soil fertility, crop yield, and yield forecast provision to improve food security. This dissertation addresses some of the most challenging aspects of using agricultural technologies to assess and increase crop yield, and to improve food security in African countries, including how to use the

technology that already exists in African countries (e.g. tablets and smart phones) to collect real-time agronomic information for forecasting yield in a timely manner, how to build soil fertility to feed the future, and how satellite images can be used in smallholder farming systems.

Objectives and Structure of the Dissertation

The overarching goal of this dissertation is to develop data-driven agricultural technologies to improve food security in the background of increasing demand from population growth and a changing climate with warming temperatures and unreliable precipitation events. This dissertation consists of an introductory chapter, four research chapters, and a concluding chapter with summaries and recommendations. Each research chapter contains a research study. Chapter 2 presents the first study where digital cloud-stored field questionnaire surveys conducted by local enumerators and a crop model's results were integrated to provide timely in-season maize yield forecasts for smallholder farms across Tanzania. This chapter has been published in *Food Security* with open access. With the field data collected for this chapter and a crop model (described in the next section), I designed a simulation experiment to address climate impacts and uncertainties in future maize production in Tanzania, which is presented in Chapter 3. Chapter 4 investigates sustainable yam cultivation in Ghana and evaluates management practices that would increase yam tuber yield and soil organic carbon based on crop modeling. This chapter is currently under review in a peer-reviewed journal. Chapter 5 addresses the opportunities and challenges in using satellite images for monitoring small fields in low-income countries.

The Systems Approach to Land Use Sustainability (SALUS) crop model was used in multiple chapters in this dissertation. In the next section of this chapter, I provide a description of

the SALUS model's mechanism for simulating crop growth and development and the associated environmental impact. I did not provide an overview of SALUS model in the research chapters (Chapter 2-5) to avoid repetition, but the SALUS model was validated for each study in the research chapters; the validation procedures and results are included in each of the individual research chapters.

A Brief Overview of the SALUS Model

The SALUS model is a process-based crop model (Basso et al., 2006). The SALUS model was adapted from the CERES (Crop Environment REsource Synthesis) model with a series of algorithmic updates for soil nutrient and soil water simulations (Basso and Ritchie, 2015; Basso et al., 2016c). The model uses daily weather information (minimum and maximum temperature, precipitation, and solar radiation), soil layer properties (e.g. clay and silt content, bulk density, and organic carbon content), management (e.g. planting dates, planting density, and fertilization rates) and crop parameters as inputs. It simulates agronomic outputs (e.g. yield and leaf area index) and environmental impact (e.g. SOC, nitrate leaching, and water fluxes) on a daily basis. The three interconnected modules — crop growth and development, soil nutrient dynamics, and water balance — are the main components of the model. The SALUS model uses two modes to simulate crop growth: simple and complex. SALUS simple mode is similar to the Environmental Policy Integrated Climate (EPIC) model. The simple mode was designed to capture crop growth using the thermal time-leaf area index curve with considerations for nutrient cycle and water balance. In SALUS simple mode, daily potential biomass accumulation is first calculated based on leaf area index, solar radiation, radiation use efficiency, and CO_2 (Dzotsi et al., 2013). The potential biomass is then reduced based on abiotic stress, including N, P, water deficiency and

heat stress, to calculate final daily biomass accumulation. Nutrient and water stress factors in the SALUS model are quantified as supply:demand ratio. A value of 1 indicates no nutrient or water shortage and therefore no stress, and a value of 0 means severe stress with no supply to meet the demand (Liu and Basso, 2017a). Lastly, daily biomass accumulation is partitioned into harvestable biomass yield (e.g. grain yield and tuber yield) and non-harvestable biomass (e.g. stalks and leaves for maize, and vines and leaves for yam) using harvest index (Dzotsi et al., 2013). Regardless of SALUS mode, the model determines crop development based on minimum and maximum temperature, base temperature, and optimal temperature (Basso et al., 2006; Dzotsi et al., 2013). The SALUS model does not have an explicit module to account for pest, weeds, and diseases.

The nutrient cycle module in SALUS was derived from the CENTURY model with modifications. Three soil carbon pools with different turnout rates are considered in the SALUS model, active, slow, and passive. The sizes of the three pools are initialized based on procedures in Basso et al. (2011). N cycling processes, including soil organic matter decomposition, immobilization, mineralization, and transformation to gaseous N are also simulated. Immobilization and mineralization of carbon and N is determined based on the C:N ratio (Basso and Ritchie, 2015; Liu and Basso, 2017a). SALUS considers three inorganic P pools (labile, active, and stable) and two organic P pools (active and stable) for P cycling. Algorithms and procedures for the P cycle have been described in Daroub et al. (2003).

The water balance module considers precipitation, runoff, evaporation, transpiration, and drainage processes, and was adapted from the CERES model with revised algorithms for infiltration, drainage, evapotranspiration, and runoff (Basso et al., 2010; Syswerda et al., 2012). Algorithms for crop evapotranspiration were based on Ritchie's equations (Ritchie, 1972). Crop

transpiration is calculated from potential evapotranspiration. The equation to partition transpiration from potential evapotranspiration depends on soil surface wetness and leaf area index. Algorithms used for evapotranspiration simulation in SALUS were presented in Basso and Ritchie (2012) and Ritchie and Basso (2008). Soil water infiltration, drainage, and runoff is simulated based on the time-to-pond concept and a physical-based water redistribution model, instead of the Soil Conservation Service runoff curve numbers (Basso et al., 2010). SALUS's soil water distribution algorithms have been described in Suleiman and Ritchie (2003) and Ritchie et al. (2009).

The SALUS model has been validated across various climatic environments, including humid subtropical (Albarenque et al., 2016; Liu and Basso, 2017a), Mediterranean (Cillis et al., 2018), tropical (Liu and Basso, 2017a, 2020b), and warm humid continental climates (Basso and Ritchie, 2015; Liu and Basso, 2020a). It has been applied to evaluate both cereal and non-cereal production under varying management practices and under historical and future climates. Particularly, Liu and Basso (2017b) developed a modeling framework to identify constraining factors for switchgrass cultivation across Michigan and to evaluate the effect of nitrogen fertilizer application rates on switchgrass yield enhancement across Michigan under historical and projected climate scenarios. Liu and Basso (2017a) used a crop model to quantify the impact of soil and climate on maize yield across Malawi at 0.25-degree resolution and to evaluate the effect of alternative management scenarios (e.g. maize and pigeonpea rotation, maize with pigeonpea residue incorporation, and maize with N fertilizer) on yield. In addition, the SALUS model has been evaluated to simulate yields of maize, soybean, and wheat (Basso and Ritchie, 2015; Liu and Basso, 2020a), soil water content (Basso et al., 2010; Basso and Ritchie, 2015), soil carbon (Cillis et al., 2018; Pezzuolo et al., 2017), and soil nitrate and nitrate leaching (Basso

et al., 2016b; Giola et al., 2012) in a Mediterranean climate in Italy and humid continental climate in the US.

CHAPTER 2: LINKING FIELD SURVEY WITH CROP MODELING TO FORECAST MAIZE YIELD IN SMALLHOLDER FARMERS' FIELDS IN TANZANIA

A version of this chapter appeared in the journal *Food Security* (doi:10.1007/s12571-020-01020-3)

Abstract

Short term food security issues require reliable crop forecasting data to identify the population at risk of food insecurity and quantify the anticipated food deficit. The assessment of the current early warning and crop forecasting system which was designed in mid 80's identified a number of deficiencies that have serious impact on the timeliness and reliability of the data. We developed a new method to forecast maize yield across smallholder farmers' fields in Tanzania (Morogoro, Kagera and Tanga districts) by integrating field-based survey with a process-based mechanistic crop model. The method has shown to provide acceptable forecasts (r^2 values of 0.94, 0.88 and 0.5 in Tanga, Morogoro and Kagera districts, respectively) 14-77 days prior to crop harvest across the three districts, in spite of wide range of maize growing conditions (final yields ranged from 0.2-5.9 Mg/ha). This study highlights the possibility of achieving accurate yield forecasts, and scaling up to regional levels for smallholder farming systems, where uncertainties in management conditions and field size are large.

Introduction

Crop yield forecasts provide a distribution of expected yield prior to crop harvests (Basso and Liu 2019). Knowing the expected yield of major food commodities in advance of harvesting is critical for national food security (Jayne and Rashid 2010; Stone and Meinke 2005). At a

national level, food production forecasts are used to make decisions on importing or exporting food commodities and their trading prices (Delincé 2017). Food policies regarding trading affect national food supply and food security in Africa (Wright and Cafiero 2011; Sitko et al. 2018). At a field level, food supply is determined by crop productivity. In-season crop yield forecasts provide management suggestions to optimize resource use efficiency (e.g. nitrogen fertilizer) and to achieve yield potential at a field level (Raun et al. 2005; Zinyengere et al. 2011).

Many countries have institutional infrastructure for operational crop yield forecasts for strategic planning. Government agencies are involved in providing information about field conditions, crop status and weather conditions to release multiple stage yield forecasts before planting, during the growing season, and prior to harvest (Gennari and Fonteneau 2016). The assessment of the current early warning and crop forecasting system which was designed in mid 80's identified a number of deficiencies that have serious impact on the timeliness and reliability of the data (Basso and Liu 2019; Gennari and Fonteneau 2016; Luo et al. 2011).

There are three major approaches to forecasting crop yield: expert-based assessments (e.g. interviews and field surveys), statistical models, and process-based models (Basso and Liu 2019). Interviews with farmers can provide subjective expectations on end-of-season yield (Nandram et al. 2014; Pease et al. 1993). Field surveys with crop cutting provide objective yield estimates prior to harvesting. Statistical models apply different techniques (regression, Bayesian approaches, machine learning techniques) to relate historical yield records to historical within-season agrometeorological variables, variables derived from remotely sensed vegetation indices and/or crop model outputs to predict yield based on the growing-season information (Johnson 2014; Lobell et al. 2015). Crop models produce not only end-of-season yield but also yield distributions based on crop genotypes, soil conditions, typical management practices, and in-

season weather based on historical climate or weather forecasts, or by assimilating remotely sensed information (Arkin et al. 1980; Jones et al. 2017; Kadaja et al. 2009; Reynolds et al. 2000).

Despite the extensive studies on yield forecasting methodology, most of the work has been done for developed nations where fields are likely to be large with one single crop per growing season, while only a small fraction of the literature has focused on yield forecasting methods for smallholder farming systems, either pure stands or intercropping (Basso and Liu 2019). Much work done so far in yield prediction has explored the use of statistical agrometeorological models, where yield is forecasted based on in-season agronomic (e.g. leaf area index, fertilizer use, and planting date) and meteorological data, either from observations or derived from satellites (Basso and Liu 2019; Choularton et al. 2019; Coughlan de Perez et al. 2019). Because the accuracy is constrained by the ranges of agrometeorological conditions that were included in the model development, the scalability of the statistical models to different years, to other regions and to other crops is limited (Katz 1977). A few researchers have used process-based model with previous seasonal weather data to provide yield forecasts for sorghum in Burkina Faso (Mishra et al. 2008), and within-season satellite-derived rainfall estimate for maize yield forecasting in South Africa and Kenya (Lourens and De Jager 1997; Reynolds et al. 2000).

A recent advance in the statistical models for yield forecasts for smallholder farmers was the development of regression models with growing-season weather, remotely sensed vegetation indices, and crop model outputs and their applications to estimate final yield (Burke and Lobell 2017; Lobell et al. 2015). This approach relies on growing-season weather information (e.g. rainfall in the last months of the growing season) or vegetation indices (e.g. peak normalized difference vegetation index) and provides yield estimates before the end of the growing season.

The lead time and skills of maize yield forecasts are limited for smallholder maize cropping systems in Africa. In most studies, the yield forecast was delivered at harvest time but not during the growing season (Basso and Liu 2019). One attempt was made to forecast maize yield using agrometeorological models, which were based on both climatic variables from weather stations and vegetation indices derived from satellite imageries, at initial and vegetative stage in Kenya (Rojas 2007). The reported r-squared values from the forecasted and observed yield regression model were mostly less than 0.5 when the forecasts were made at vegetative or reproductive stage (Mkhabela et al. 2005; Rojas 2007; Schauberger et al. 2017). A few cases in Kenya and Swaziland had higher correlation between the forecasted yield, made a few months before harvest, and the final yield, with r-squared values greater than 0.7 (Mkhabela et al. 2005; Rojas 2007). Others used regression models to forecast yield in Zimbabwe and Botswana at maturity based on climatic variables (Manatsa et al. 2011; Vossen 1990) and satellite derived vegetation indices (Kuri et al. 2014; Unganai and Kogan 1998), and obtained adequate forecasting accuracy, with r-squared values over 0.8. It has also been noted that forecasting procedures, particularly when statistically based, performed much worse when applied to smallholder farming systems in Africa, compared to large farms in the US (Azzari et al. 2017; Schauberger et al. 2017).

In this study, we present a new maize yield forecasting method that provides yield forecast for governmental agencies before the crop is harvested (14-77 days prior to harvest). The objective of this paper was to develop and validate a new method to forecast maize grain yield based on the integration of field survey and crop model in three regions in the United Republic of Tanzania (Tanzania hereafter).

Materials and Methods

Context of the Research Project

Accurate and reliable crop yield forecasting data to identify the population at risk of food insecurity and quantify the anticipated food deficit is a key policy concern of the Government of Tanzania. The current forecasting system of Tanzania presents a number of deficiencies that have serious impact on the timeliness and reliability of the data.

Improvement of the crop forecasting system was one of the actions identified under Strategic Objective 3 ("rationalize statistical operations and processes, improving quality and relevance to users of agriculture statistics data") of Tanzania Agricultural Statistics Strategic Plan (ASSP) recently prepared and adopted by the government. This research project, designed to develop a new and practical method to provide accurate and timely crop yield forecasts for the Government of Tanzania, was selected under the framework of the Global Strategy to Improve Agricultural and Rural Statistics coordinated by the United Nations Food and Agriculture Organization (UN FAO TCP URT 3504).

Descriptions of the New Yield Forecasting Method

The new yield forecasting method presented in this study is based on the integration of field based surveys and the process-based crop model SALUS (Systems Approach to Land Use Sustainability) (Figure 1).

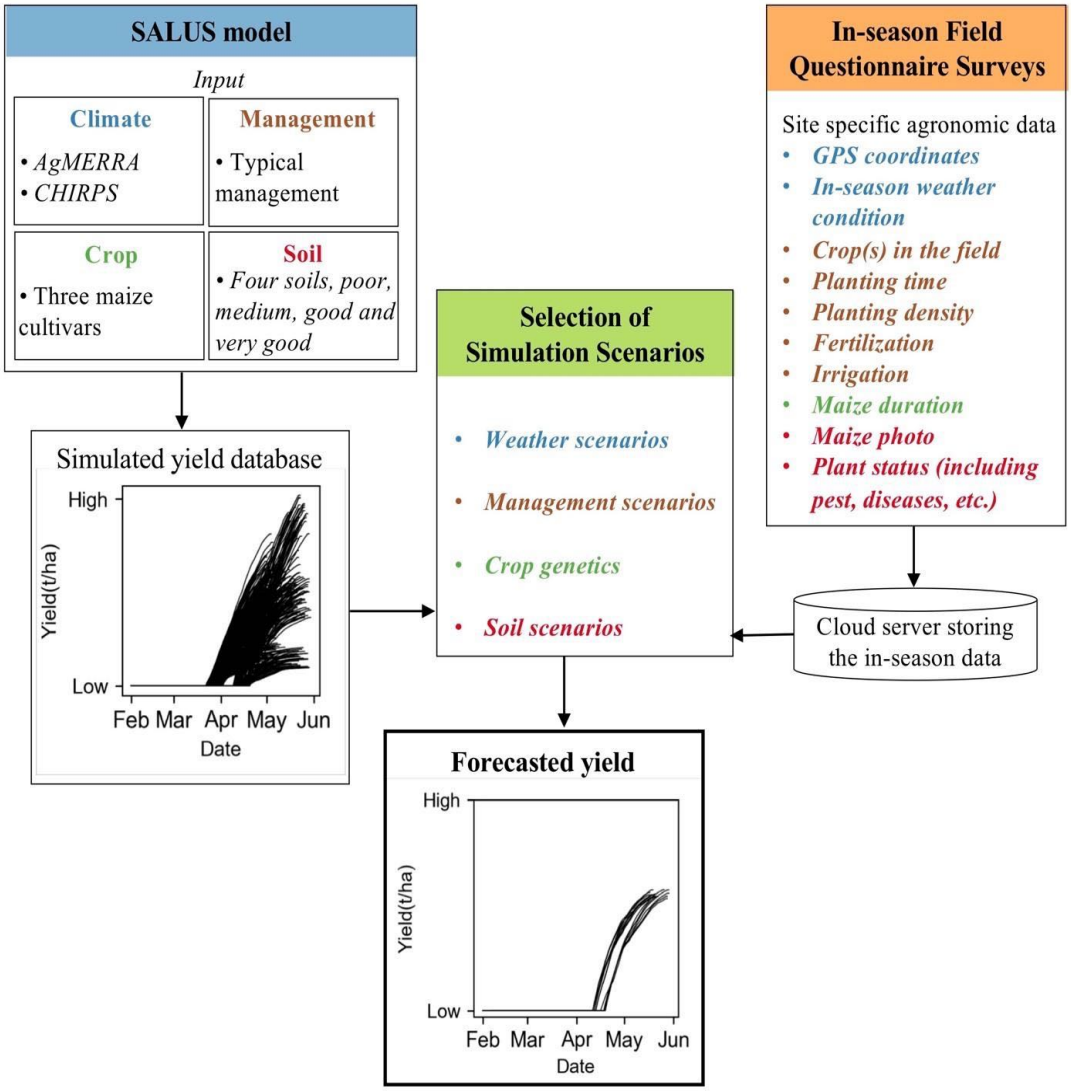

Figure 1 Overview of the new yield forecasting method based on the integration of field surveys and the SALUS crop model.

SALUS Model Execution

For this study, we tested SALUS model to evaluate its capability to reproduce interannual maize grain yield at the regional level (APPENDIX A Figure 28 and APPENDIX A Figure 29).

In this study, we ran SALUS using weather, crop, soil, and management inputs that captured possible scenarios reported by field surveys. The weather data in the study were based on the 0.25º-resolution AgMIP climate forcing based on the Modern-Era Retrospective Analysis

(AgMERRA) dataset and the 0.05º-resolution Climate Hazards Group InfraRed Precipitation with Station (CHRIPS) dataset (Funk et al. 2015; Ruane et al. 2015). We extracted daily temperature and solar radiation data in 1981-2010 from the AgMEERA, and daily precipitation data from the CHIRPS gridded dataset. We used three maize cultivars in the simulations to represent short, medium and long duration cultivars (APPENDIX A Figure 30). We included four options to describe soils: poor, medium, fertile and extremely fertile (APPENDIX A Table 6). Management practices (planting densities, N fertilizer application rates irrigation amounts) were reported by field survey information for the districts of Morogoro, Kagera, and Tanga).

Field Questionnaire Survey

Currently, the Government of Tanzania collects in season information from farmers' fields using 7 field questionnaire surveys to monitor crop conditions and forecast crop yield and production. We designed a simple questionnaire to collect field data as model input on agronomic management, weather and plant information.

Management information included data on number of plants present in selected fields at the time of the field survey, maize cultivar characteristics (short versus long duration), planting time, planting density, fertilizers and irrigation amounts and timing of applications. For the weather conditions, we asked qualitative descriptions of the current growing-season temperature and rainfall conditions when compared to historical averages (options included: hotter than normal, colder than normal, normal, drier than normal, wetter than normal). For the plant conditions at the survey date, we asked the enumerators to take photos of the maize plants grown in the fields to detect presence of diseases, weeds and insects. The questionnaire is available in APPENDIX A Table 7-8.

During the season, data collection, both interviews and field sampling, were conducted by 55

enumerators to complete the questionnaire. A quadrant design method was used to collect planting density and plant condition information. One or two experimental plots (6x6 meter) were first randomly established at each sampling field. Quadrants were then formed within the experimental plot. Planting density was based on the number of plants across the four quadrants. Two plants in each quadrant were randomly marked for the end-of-season field survey. In the end-of-season field survey, crop cut and kernel numbers were performed to estimate grain yield. Cobs in the experimental plots were harvested and weighed. The two plants previously marked within each quadrant were sampled for total number of kernels and kernel weight. The kernel information was then converted to determine the grain yield.

Yield Forecast Method

The SALUS crop model was executed using a combination of a series of soils, weather, genotypes and management practices. The method searched for the simulation scenarios that best represented the growing season conditions reported by the field survey questionnaire, and the simulated yields of the selected simulations served as one of the forecasted yields depending on the remaining weather to reach crop maturity. For each of the sampling fields, the yield forecasting algorithm used the reported coordinates to identify historical weather scenarios among the climate dataset, and then selected years in which temperature and precipitation matched the reported in-season temperature and precipitation characteristics. The algorithm first grouped historical years into three categories based on the 33.3^{th} and 66.7^{th} percentile values of average seasonal temperature in 1981-2010. Years where the in-season average temperature was less than 33.3^{th} percentile of the average temperature in 1981-2010 were categorized as colder than normal; years where the in-season average temperature was more than 66.7^{th} percentile of the average temperature in 1981-2010 were categorized as hotter than normal; years where the

in-season average temperature was between the 33.3th and 66.7th values was normal. Similarly, years were also grouped into drier than normal, normal and wetter than normal categories based on the 33.3th and 66.7th percentile values of total seasonal precipitation in 1981-2010. The algorithm then selected weather series where the temperature and precipitation categories matched with the reported in-season weather characteristics. In the cases where the no historical record was found to match with both in-season temperature and precipitation characteristics, the algorithm prioritized matching with the reported precipitation condition. The algorithm narrowed management scenarios based on the reported planted time, plant densities, and fertilizer and irrigation applied. The yield forecasting algorithm used the reported maize duration to exclude simulation scenarios in which the duration did not match the reported value. The yield forecasting algorithm used the overall evaluation of the field to select soil used in the simulations. The reported stress level due to water and nitrogen deficit, together with the photos taken during the survey determined the overall field condition. We developed a protocol to evaluate the overall condition of maize fields based on the photos and reported stress level. When a field had maize with healthy dark green leaves and relatively thick stalks, it was categorized as extremely good condition. When the plant was mostly dark green but under minor stress, it was categorized as in good condition. Medium condition indicated plants with light green leaves and under nitrogen and/or water deficit stress. Poor condition indicated short plants with yellow-green leaves and thin stalks and were under severe stress level. We used the photo to cross check the reported stress level due to water and nitrogen deficit and biotic stressors (i.e. weeds, pests and diseases). The selected simulation scenarios (at least one simulation runs) contained combinations of the selected weather, management, crop and soil scenarios, which resulted from the yield forecasting algorithm. Lastly, the simulated yields of the selected

simulations were adjusted based on the severity of insects, weeds and diseases. We applied a 15% reduction to the simulated yield when the severity of insects, weeds or diseases was reported minor, and a 30% reduction when the severity was reported major (Tollenaar et al. 1994). The simulated yields, with adjustment for weeds and insects if reported, were the forecasted yield for each sampling field.

Study Sites and Date Collection

The method was applied to three districts in Tanzania, Morogoro, Kagera and Tanga. The study sites were selected by officials from the Government of Tanzania and the UN FAO (Figure 2a). The three districts have equatorial climate but have distinct agro-climatic characteristics (Kottek et al. 2006). Kagera is located in northern Tanzania and has bimodal rainfall pattern, where short rain starts in October and ends in December, and long rains start in March and end in May. Tanga is located in northeastern Tanzania and also has bimodal rainfall pattern. Maize is widely cultivated in Tanga whereas banana is an important crop in Kagera (Smale and Tushemereirwe 2007). Morogoro is located in central Tanzania transitioning between bimodal and unimodal rainfall (unimodal rainfall occurs between November and May) (Paavola 2008).

A total of 92 sampling fields across three districts were determined. Specifically, 28 sampling fields were located across Morogoro, 39 across Kagera and 25 across Tanga. The enumerators conducted within-season field questionnaire surveys spanning from late April to June 2017 for Morogoro, from mid-January to the end of February for Kagera, and from late January to late March for Tanga. The majority of the surveys were completed by end of May in Morogoro, and by early January in Kagera and Tanga. Of the predetermined sampling fields, 17 fields in Morogoro had maize plant that were not mature during the in-season field survey, as well as 24 fields in Kagera and 21 fields in Tanga (Figure 2b-d). For the remaining fields, maize either

reached maturity (11 fields in Morogoro, 13 in Kagera and 4 in Tanga) or the field survey was incomplete (two fields in Kagera), and thus they were not included in the analysis. The 62 sampling fields were across 55 households in the three districts.

The questionnaire we developed was coded in the Survey Solutions application in both English and Swahili. Survey Solutions is a computer-assisted personal interviewing software developed by the World Bank. The trained enumerators administered the field questionnaire survey using tablets with the questionnaire coded in the Survey Solutions application. The enumerators recorded the geographic location and surveyed the physical characteristics of the within-season plant (including planting density, stress level due to N, drought, weeds, pests and diseases) condition. Other in-season information (including weather characteristics and maize cultivar, sowing time, irrigation and fertilization levels) were from enumerators' interviews with the farmers or farm workers. The complete survey was synchronized to the cloud storage. We processed the within-season information immediately after we received it through the cloud storage and provided the maize yield forecast for each of the sampling fields. We provided yield forecasts ranging from 14 to 77 days prior to harvest. The 25^{th} and 75^{th} percentile of the forecasting lead time was 30 and 55 days before harvest, respectively.

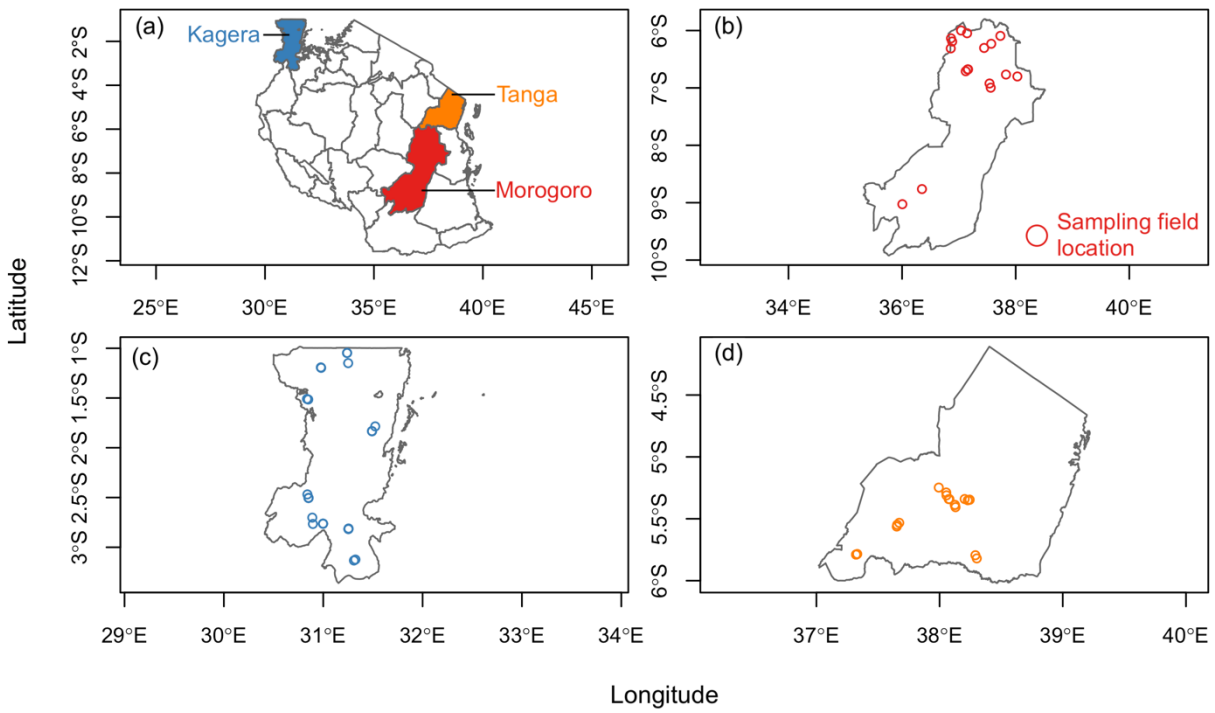

Figure 2 Sampling locations across Morogoro, Kagera and Tanga districts of Tanzania. (a) Map of Tanzania and the three districts, (b) spatial distribution of sampling fields in Morogoro, (c) spatial distribution of sampling fields in Kagera and (d) spatial distribution of sampling fields in Tanga.

Accuracy Assessment

We assessed the accuracy of the maize yield forecasts submitted before harvest based on three accuracy indicators. We first regressed the forecasted versus the observed yield and used the coefficients of determination (r^2) of the linear model to evaluate the overall agreement between the yield forecasts and yield observations. We then calculated the root mean square of deviation (RMSD) based on the Equation (1) to assess the deviation of the forecasted values from the observed ones. The RMSD value is sensitive to extreme values. Lastly, we calculated the Mean Absolute Error (MAE) (Equation (2)) to assess the accuracy of our methodology. The MAE value represents the overall error. It is a more direct representation of model biases and is less sensitive to extreme values compared to the RMSD (van der Velde and Nisini, 2019; Willmott and Matsuura, 2005).

$$RMSD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{i=N}(F_i - O_i)^2}$$ ……..……..…….. Equation (1),

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|F_i - O_i\right|$$ ……...……..……..……Equation (2),

where N is the total number of sampling fields, i is each sampling field, F is forecasted yield and O is observed yield.

We showed the accuracy of the yield forecasts for all sampled maize fields, regardless of the maize development stages during initial visit, and the accuracy for sampled fields where maize was present during the growing season.

Results

Descriptive Statistics of within-Season Data Collection

More than 90% of the sampling fields across Morogoro, Kagera and Tanga were smallholder (less than 2 ha farm area) farming systems. The size of the fields ranged from 0.056 ha to 7.49 ha with a median value of 0.83 ha in Morogoro. For the other two districts, all sampling fields were under 2 ha. The field size was 0.19-1.30 ha with a median value of 0.47 ha for Kagera and was 0.02-1.92 ha with a median value of 0.71 ha for Tanga (APPENDIX A Figure 31).

Across the three districts, maize was at early to mid grain filling stage for all sampling fields except for one field in Morogoro, where maize was at vegetative stage during the field survey. The agronomic and climatic conditions varied across the sampling sites in the three districts.

In Morogoro, more than 75% of the fields were monoculture maize and four fields had maize intercropped with either pigeonpea or field peas. Long-duration maize cultivars were reported for 12 fields and the remaining five fields had short-duration maize. The reported sowing time was between early February and early March for most of the fields, and in mid March for three fields.

Maize planting density across the 17 sampling fields ranged from 1.0 plants/m^2 to 5.7 plants/m^2 with an average value of 2.8 plants/m^2. Most of the fields was unfertilized and rainfed. Irrigation was reported for two fields and manure application was reported for one field. No synthetic fertilizer was reported across the fields in Morogoro (Figure 3). Water deficit was not reported. No N deficit was reported for 9 fields, minor N deficit for 7 fields and severe N deficit for one field. During the in-season field survey, weeds were not present for 8 fields, whereas the other 8 fields experienced minor weed problems, and one field had severe weed issues. Insects were not present in 12 sampling fields, four fields had a minor insect problem and one field had a severe insect problem. Only three fields were reported to have minor disease issues and the remaining majority of fields did not have disease problems. Based on photos taken during the in-season field survey, one field was assessed in extremely good condition with dense plants, healthy green leaves and relatively thick stalks, eight fields were in good condition, four fields were in medium condition with yellow spots on green leaves and relatively thin stalks, and four fields were in poor condition with short plants, very thin stalks and/or unhealthy leaves (Figure 4).

In Kagera, slightly more than half of the sampling fields (14 out of 24 sampling fields) were pure maize stands and the remaining fields were maize intercropped with banana (four fields), banana and beans (one field), banana and cassava (one field), beans (two fields), or cassava and beans (two fields). More than 60% of the sampling fields (15 fields) had long-duration maize cultivars and the remaining fields had short-duration cultivars. Early October or late August were the predominant sowing times. The sowing time was in early, mid or late September for a total of 7 fields. Maize plant density was low across the sampling fields, ranging from 0.3 plants/m^2 to 3.3 plants/m^2 with an average value of 1.3 plants/m^2. Rainfed maize was reported for 23 out of 24 fields. A few fields (four fields) had manure applications by the survey date and none had

synthetic fertilizer input. A majority of the maize fields had no water deficit whereas three fields were had minor water deficit conditions (Figure 4). Minor N deficit was reported for a majority of the fields (22 fields) and only two fields had no N deficit conditions. Half or more of the fields were reported to have minor weed, insect and disease problems. Weeds, insects and diseases were not reported in 7, 5 and 11 of the sampling fields in each district, respectively. Severe weed, insect and disease problems were reported for one or two fields. We assessed that 10 fields were in good condition, 9 were in medium condition and five were in poor condition (Figure 4).

Maize was grown in pure stands across the 21 sampled fields in the Tanga district. Of the sampled fields, 14 fields were sown with a short-duration cultivar and seven had long-duration cultivar maize. Maize was sown in early October for more than 90% of the fields and was planted in late September for two fields. Maize plant density ranged from 0.4 plants/m^2 to 3.6 plants/m^2, averaging 1.9 plants/m^2. A majority of the maize fields were rainfed and unfertilized. Irrigation was reported for three fields. Manure application was reported for two fields (Figure 3). None of the sampled fields had synthetic fertilizer input. Maize experienced minor water deficit conditions in more than half of the fields, while severe water deficit stress was reported for four fields. A majority of maize fields experienced minor N deficit and two fields had adequate N supply. Most maize fields had minor or severe weed, insect and disease problems. Overall, three fields were in extremely good condition, four were in good condition, 9 were in medium condition and five were in poor condition (Figure 4).

Regarding in-season climatic characteristics, a majority of the respondents in Morogoro reported hotter than normal and drier than normal condition, one respondent reported colder and drier than normal conditions, and the other three reported normal temperature and rainfall conditions. By contrast, most respondents in Kagera district reported average rainfall and

temperature conditions when compared to the historical norm, three reported wetter and colder than the norm, and the other three reported wetter but hotter conditions than the norm. For Tanga, 13 respondents reported drier and hotter than normal conditions, 7 reported average rainfall and temperature conditions, and one reported drier and colder than the norm (Figure 3).

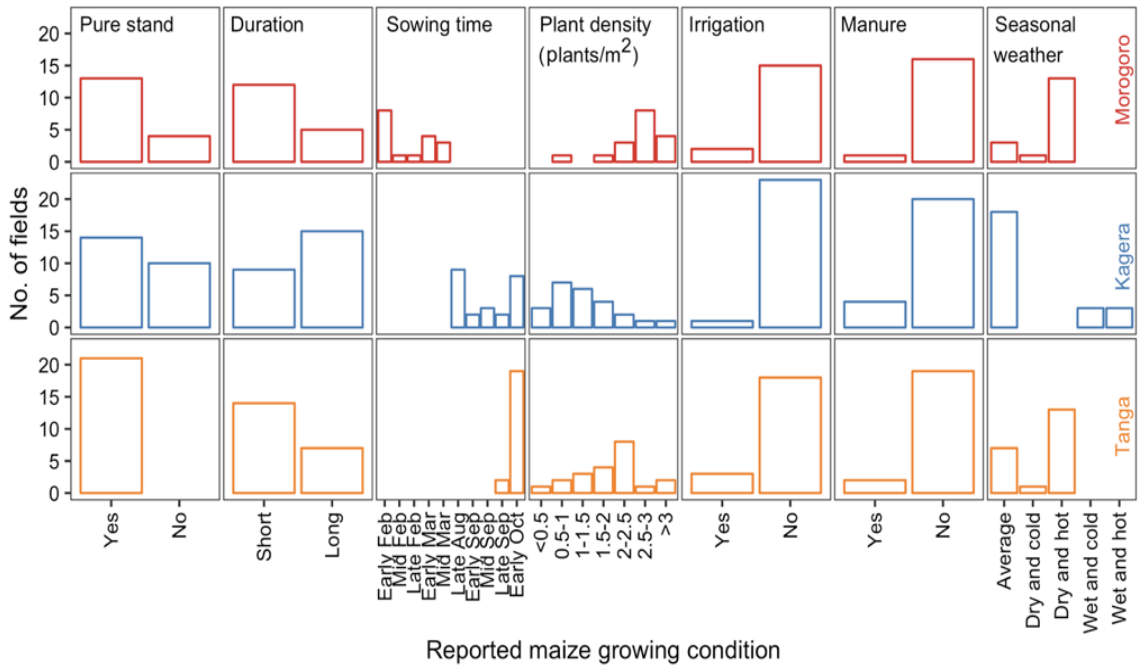

Figure 3 Reported maize growing conditions, including pure crop stands versus intercropping, maize duration, sowing time, plant density, irrigation and manure use and growing season weather characteristics across the three districts.

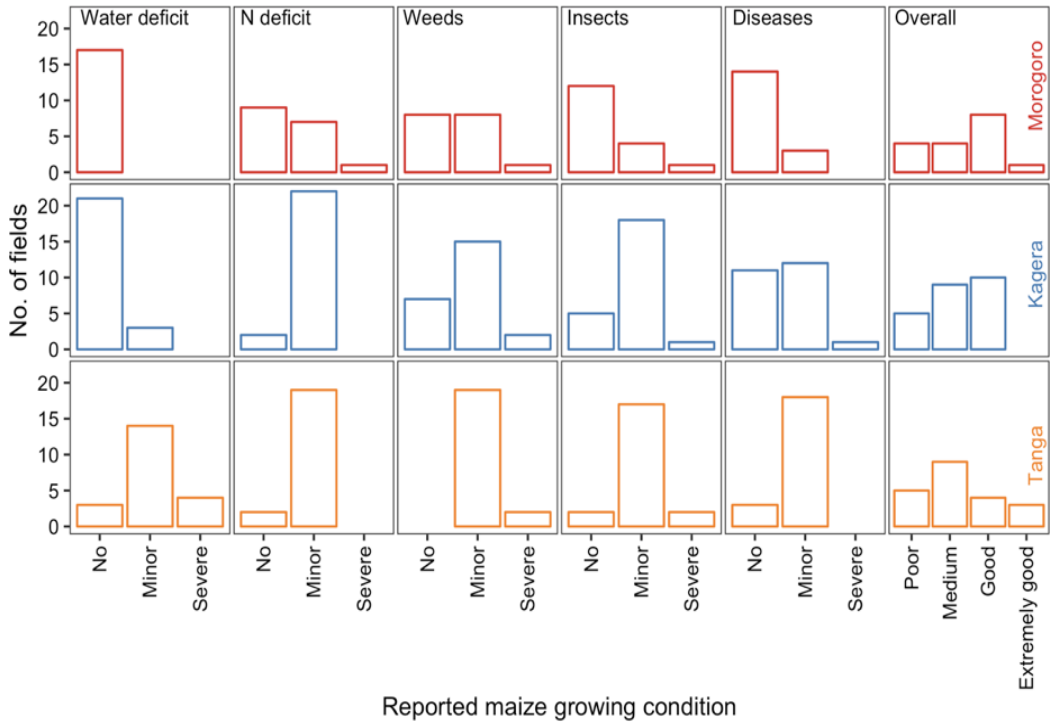

Figure 4 Maize status, including water and N deficit, weed, insect and disease presence, and overall plant condition based on photos taken during in-season survey across the three districts.

Maize Yield Forecasts

Our proposed method was able to accurately forecast maize yield before the harvest across the three districts under varying conditions. Maize yield ranged from 0.74 to 5.63 Mg/ha with an average value of 2.36 Mg/ha and standard deviation of 1.2 Mg/ha in Morogoro. The forecasted yield captured the variations in the reported final yield, with an r^2 value of 0.88. The RMSD value between the forecasted yield and the reported yield was 0.47 Mg/ha and the MAE value was 0.36 Mg/ha (Figure 5a).

Maize yield in Kagera was low, 0.19-1.94 Mg/ha with an average value of 0.94 Mg/ha and standard deviation of 0.51 Mg/ha. Using our proposed method, we were able to closely forecast the final yield for most fields (r^2= 0.5). The RMSD was 0.38 Mg/ha and the MAE was 0.25 Mg/ha (Figure 5b).

For Tanga, where all maize fields were monoculture, maize yield ranged from 0.28 to 5.84 Mg/ha, with an average value of 2.03 Mg/ha and standard deviation of 1.59 Mg/ha. The forecasted yield closely matched with the reported yield with r^2 value of 0.94. The RMSD value between the forecasted and the reported yield was 0.43 Mg/ha and the MAE value was 0.32 Mg/ha (Figure 5c).

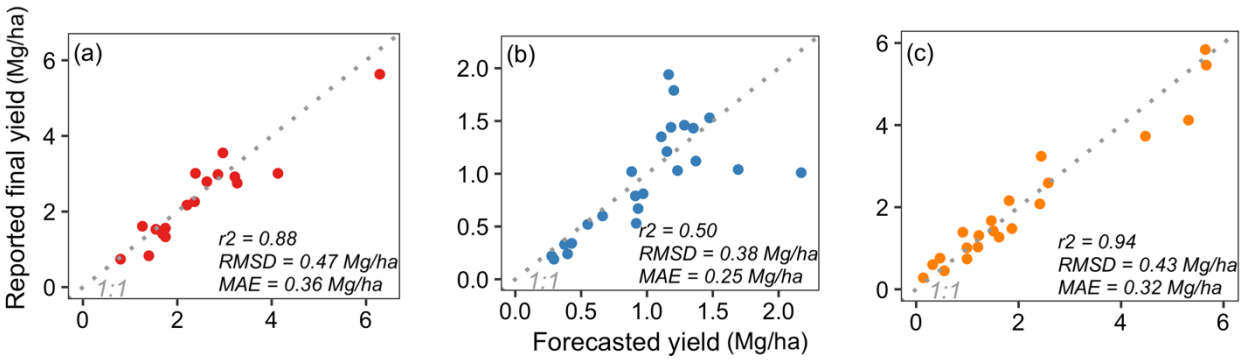

Figure 5 Comparisons between the forecasted yield and reported final yield across (a) Morogoro, (b) Kagera and (c) Tanga (note that the ranges for both axes in a-c differ).

Discussion

Forecasting grain yield before harvest for smallholder farming systems has been a major challenge for scientists and government officials working on this important topic. Crop models supplied with daily weather observation till the forecasting date combined with historical observations for the remaining growing season, weather analog or seasonal weather forecast have been used to generate yield forecasts (Hansen and Indeje 2004). Due to lack of extensive weather station network in Africa, weather observations are limited and seasonal climate forecasts have low skills (Sheffield et al. 2014; Singh et al. 2018). Quality real-time weather data was not publicly accessible for our studied site. Instead of relying on real-time weather data to simulate in-season crop growth and development, we used descriptive in-season weather characteristics from the field questionnaire survey to select analog years in long-term reanalysis

climate datasets (AgMERRA and CHIRPS). This was the first study, to the authors' knowledge, that linked field surveys with a crop model to forecast crop yield. The other innovation featured in our study was the use of digital photos, taken with inexpensive and widely available mobile computer tablets. The photos were used to assess growing season plant conditions, including nutrient and water deficit levels, and weed, insect and disease issues. These photos provided valuable mid-season information for experts to cross validate the answers reported by the enumerators in the field survey questionnaire. We also implemented a simple reduction factor to consider the effect of biotic stress (weeds, diseases and pests) on yield, which was not simulated by crop models but prevalent in smallholder farms.

We demonstrated that it was feasible to make accurate yield forecasts by collecting both subjective and objective in-season yield assessment as well as running the crop model. The use of subjective evaluation of field condition in yield forecasts has been employed by the US Department of Agriculture and has been documented in the literature (Pease et al. 1993 and USDA 2012). Our methodology produced accurate yield forecasts for smallholder farming systems (less than 2 ha) across three districts in Tanzania one to two months before harvest (r^2 values of 0.94, 0.88 and 0.5 in Tanga, Morogoro and Kagera districts, respectively). The accuracy of our method was much higher than most of the yield-forecasting studies for Africa. The r^2 values between the forecasted yield and final yield were mostly under 0.5 when forecasts were made during maize vegetative to reproductive stages (Abo-Shetaia et al. 2005; Mkhabela et al. 2005; Schauberger et al. 2017) with an exception of Unganai and Kogan (1998), where regional maize yield in Zimbabwe was accurately forecasted ($r^2 > 0.9$) by the peaked Temperature Condition Index and Vegetation Condition Index derived from the Advanced Very High Resolution Radiometer satellite. There were a few studies reported comparably accurate

yield forecasts at maturity ($r^2 > 0.8$) but our method offered accurate forecasts 1-2 months before harvest (Manatsa et al. 2011; Rojas 2007; Vossen 1990). Basso and Liu (2019) has provided a comprehensive review on crop yield forecasting methods and their accuracy.

Despite that statistical-based models could provide accurate crop yield forecasts for smallholder fields in African countries (e.g. Manatsa et al. 2011 and Rojas 2007), statistical yield forecasting model may not be applicable to growing conditions that are beyond the model development boundary. Crop models have an advantage over statistical models regarding capabilities of representing crop growth and development under climate change (Lobell et al. 2017; Jones et al. 2017). Our approach of integrating crop modeling and surveys provides a framework to develop new yield forecasting methodology for other sites and other crops during other years. Our approach can be applied to forecasting maize yield in Tanzania under the long-term climate change conditions as well. To apply our method to maize yield forecasts in another country or region, one needs to identify proper soil information and to determine historical or in-season climate to cover the growth conditions. In this study, we focused on maize, the dominant staple food crop in Tanzania. Our yield forecasting procedure, nonetheless, can be applied to other crops, as long as the crops of interest can be simulated by the crop model. Regardless of crops of interest, the field questionnaire survey should include plant density, within-season information about weather, crop growing status (including stress levels due to weeds, pests, diseases, and N and water deficiency), and soil information.

Our approach has a few limitations. First, it was constrained by the trained enumerators' ability to conduct field questionnaire surveys. While enumerator availability may not be the constraining resources in developing countries, field campaigns can be labor intensive and time consuming. Second, the yield forecast product resulted from this study was subject to the quality

of the questionnaire response. Our algorithm relied on the subjective evaluation of the weed, disease and pest presence and the overall plant status assessment. In addition, AgMERRA weather dataset, which provided temperature and solar radiation input for the yield forecasting algorithm, had a limited temporal coverage, from 1980 to 2010. AgMERRA dataset may not be adequate to forecast crop yield with the rapid changing climate since the weather analog assumption will be violated as we are moving to a new climate regime. National Aeronautics and Space Administration Prediction Of Worldwide Energy Resource (NASA POWER) provides daily weather data since 1997 (https://power.larc.nasa.gov). The NASA POWER dataset may be an alternative to AgMERRA, though the POWER dataset has a coarser spatial resolution (0.5 arc degree). There are a few sources of uncertainties in our study. Though the soil and maize cultivar information we used for crop modeling and for creating yield database before the field questionnaire surveys adequately represented soils and maize cultivars in Tanzania, we did not have on-site soil data by depths across our study sites or the genetic parameters of cultivars grown in the sampling fields. There were uncertainties in soil and plant parameters of the SALUS model. The other uncertainty was the evaluation of in-season photos taken by enumerators. Different interpretation of the photos can lead to different forecasted maize yields. Another source of uncertainty was the final grain yield. Due to the logistics of the field survey, we asked for kernels number and cob weight but we did not shell maize cobs to weigh the kernel. We found inconsistencies between estimated yield based on kernel number and estimated yield derived from cob weight and planting density across the sampling fields, perhaps due to precipitation event before conducting the final field survey and moisture contribution to the cob weight. Due to missing cob weight and the quality of the cob weight, final grain yield was estimated based on total kernel number.

To overcome the limitations of field questionnaire survey, high resolution commercial remote sensing imageries that are increasingly available to research use may provide within-season vegetation status and information about management practices (e.g. planting date). Two research advances are needed to achieve accurate yield forecasts at the regional level using the proposed framework. First, high resolution cultivated crop maps are required to scale up our proposed framework. Second, vegetation status, planting date and plants density needs to be extracted from high resolution satellite imageries, such as PlanetScope and SkySat (https://www.planet.com).

Conclusions

We have presented a new method that integrates within-season field survey and crop modeling to forecast yield for smallholder farming system. We applied our proposed method to forecast maize yield at field scale across three districts in Tanzania, where maize was cultivated under different planting densities, varied intercropping plants and distinct growing-season weather conditions. The results showed that we achieved accurate yield forecasts across diverse maize fields. This study provided the most accurate field-level yield forecasting method for smallholder farming systems in Tanzania to date, which is a critical piece of information toward understanding areas within regions affected by food shortages or overproduction, leading to more informed decision by government officials.

Acknowledgements

This study was supported by the United Nation Food and Agriculture Organization (UNFAO) and the Government of Tanzania (GoT). The authors thank Carola Fabi, and Eloi Ouedraogo ,

Silvia Tireweshobwa of the UNFAO, Titus Miwsomba, Festo Mwemutsi from the GoT National Bureau of Statistics, Wilson Katunzi and Valerian Vitalis from the GoT Ministry of Agriculture. The authors were grateful for technical assistant from UNFAO's Michael Rahija in accessing survey data in the cloud server. Lastly, the authors expressed gratitude to enumerators for collecting the field survey data.

CHAPTER 3: EVALUATING CLIMATE CHANGE IMPACT ON YIELD, SOIL ORGANIC CARBON, AND SOIL NITRATE OF MAIZE-BASED SMALLHOLDER SYSTEMS IN TANZANIA

Abstract

Smallholder crop yields are vulnerable to climate change. Recent advances in regional climate modeling in Africa allows for assessment of the uncertainty of climate impacts on smallholder agriculture. The objectives of this study were to quantify uncertainties in the projected climate for the next 30 years (2020-2049) and to evaluate potential climate impacts on yield, soil organic carbon (SOC), and soil inorganic nitrogen (N) of maize-based systems in smallholder fields in Tanzania. I analyzed four climatic variables, average minimum and maximum temperature, precipitation, and solar radiation, using bias-corrected climate output from 18 regional climate models in the Coordinated Regional Climate Downscaling Experiment (CORDEX) program. I simulated climate impacts using the Systems Approach to Land Use Sustainability (SALUS) model at 60 sites in Morogoro, Kagera and Tanga regions in Tanzania. Management, soil, and plant inputs in the SALUS model were derived from field questionnaire surveys conducted in 2017-2018 and described in Chapter 2. Climate input was based on (i) daily weather from CORDEX's 18 climate models and (ii) a delta-method projected climate with historical climatic variables changed to match the average projected changes across the 18 CORDEX climate model simulations. The 18 regional climate models in general produced agreement on changes in average daily minimum and maximum temperature and solar radiation (coefficient of variation, CV, was within 6%) for the study sites in the three districts under Representative Concentration Pathway (RCP) 4.5 and RCP 8.5 scenarios. But sizable variations

in the projected total rainfall were observed across the 18 climate models (CV 27-55%). With the 18 climate models, the projected grain yield varied considerably (CV 3-21%) under the two RCPs. Substantial projected soil N variability was simulated by the 18 climate models (CV 31-117%). SOC was the least affected by uncertainty in climate change (CV within 4%). When using the delta-method projected climate, the simulated climate impact on soil N and SOC was close to the averaged simulated impact from the 18 climate models for most cases. The adverse impact on grain yield was smaller when using the delta method, compared to the averaged changes in yield across the climate models.

Introduction

Smallholder farms in African countries, often family-operated and resource-constrained, are particularly vulnerable to climate change (Lowder et al., 2016; Michler et al., 2019). The projected warming temperatures are less favorable for cereal crop development and growth. The other component of climate change is increased variability with more erratic rain events. Changes in variance of variables can be more impactful than changes in their means (Liu and Basso 2020a). Compared to historical climate, temperature in Africa has been projected to increase with high confidence, but the degree and direction of change in rainfall is much less certain (Masson-Delmotte et al., 2018). It is critical to evaluate the implications of climate variability of smallholder fields in Africa in the context of climate change.

Tanzania, located in East Africa, has experienced warming temperature and shifts in rainfall amounts and patterns. Yield of maize, a major staple food in the country, has been impacted by the changing climate (Kahsay and Hansen, 2016; Wainwright et al., 2019). Statistical and process-based crop models (e.g. DSSAT and APSIM) have been used to evaluate maize yield

response to climate change at field and regional scales (Mourice et al., 2017; Msongaleli et al., 2015; Rowhani et al., 2011; Tesfaye et al., 2015 and more recently Falconier et al., in press). Recently, crop modeling communities have focused on the effect of uncertainties of climate change and crop models on crop yield, with a consideration of the low-input systems in Africa (Bassu et al., 2014; Tao et al., 2018, Falconier et al., in press). But few studies have emphasized climate impacts on soil carbon and nitrogen (Basso et al., 2018a).

Long-term climate projections for Africa from regional climate models create an opportunity to evaluate climate change impact on smallholder agricultural systems in resources-limited environments. The Coordinated Regional Climate Downscaling Experiment (CORDEX) framework provides climate series driven by different global climate models (at about 2° resolution) combined with a regional climate model at various spatial (e.g. 0.44° and 0.22°) and temporal (e.g. daily and monthly) resolution for various terrestrial land masses across the globe under different greenhouse gas emission scenarios, the Representative Concentration Pathways (RCP) (Giorgi, 2019; Giorgi et al., 2009). A few studies have evaluated maize yield change under projected climate compared to historical climate in eastern-central and southern Tanzania by using crop models with projected climate from a fraction of the CORDEX regional climate models (Luhunga et al., 2017; Luhunga, 2017).

It is well known that an ensemble climate projected by multiple climate models is more accurate than that from a single model (DelSole et al., 2014; Ehsan et al., 2017). This implies a need to include multiple climate models in climate impact assessment. Recent crop modeling research has adopted the multimodel ensemble approach to assess climate impact on agriculture, with a focus on yield, using different crop models under a gradient of changes in average precipitation, temperature, and CO_2. The mean or the median from crop multimodel simulations

has been used as a base to evaluate climate impact (e.g. Asseng et al., 2017; Wallach et al., 2018). The other method for assessing the changes in the projected climate is to evaluate changes in the mean using multiple climate models (Hemer et al., 2013). Two questions arise from the crop and the climate multimodel ensemble approaches: how projected climate change impacts are distributed when using multiple climate-model outputs in a crop model, and how climate impact assessment differs when using the averaged value simulated with each member climate model, versus one simulated value with one single projected climate based on average changes in climatic variables. This paper aimed to address these two questions. The objectives of this research were to quantify uncertainties in the projected climate in the next 30 years (2020-2049) in Tanzania across the 18 regional climate models in the CORDEX; and to evaluate uncertainties in climate impact on yield, soil organic carbon (SOC), and soil inorganic nitrogen (N) of maize-based systems in smallholder fields in Tanzania.

Materials and Methods

SALUS Model Validation

I used field surveys from our previous work and field experiment observations from the literature to verify the SALUS model's ability to represent maize cultivation, including grain yield and N dynamics in Tanzania. I first compared the simulated maize yield to the observed yield in 62 fields across three districts in Tanzania (Morogoro, Kagera, and Tanga) from a field campaign in the long and short rain seasons in 2017-2018 (Liu and Basso 2020b). I then compared simulated yield and plant N uptake to the observations from a field experiment under four treatments at two sites in Tanzania, reported in Zheng et al. (2018) and Zheng et al. (2019). The four treatments included unfertilized, 50 kgN/ha fertilizer added, 100 kgN/ha fertilizer

added, and maize residue incorporated with 50 kgN/ha fertilizer applied. The experimental sites were in Iringa with low soil fertility and drier climate, and Mbeya with fertile soil and wetter climate. To test SALUS's ability to represent the maize cultivation reported by the aforementioned two papers in Tanzania, I included a total of 28 observed annual grain yield observations at the two sites under various treatments in four growing seasons in 2013/2014-2016/2017, and 12 observed plant N uptake observations, based on destructive crop sampling, followed by temperature combustion and subsequent gas analysis, at the two sites under three N input rates for two growing seasons in 2013/2014-2014/2015.

For SALUS validation across 62 fields in Tanzania, I used daily temperature and radiation records for 2017-2018 derived from the National Aeronautics and Space Administration Prediction of Worldwide Energy Resources (NASA POWER, downloaded from https://power.larc.nasa.gov) and precipitation records from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) (Funk et al., 2015). I used the soil, crop, and management information from the field survey as inputs for model validation. Details of the field survey were described in Liu and Basso (2020b). Similarly, I extracted daily weather from NASA POWER and CHIRPS, and used the reported soil and management in Zheng et al. (2018) and Zheng et al. (2019) as SALUS inputs to validate SALUS simulated grain yield and plant N uptake.

Input of Climate Impact Simulations

I obtained daily climatic variables of 18 climate scenarios from the CORDEX program. The 18 climate scenarios resulted from different combinations of global circulation models, regional climate models, and ensemble members (r1i1p1, r12i1p1 and r3i1p1). Climate scenarios in the CORDEX have been reported to reproduce the annual cycle, mean and, inter-annual variability of temperature and rainfall reasonably well for Tanzania (Luhunga et al., 2016). I first

downloaded the output of the 18 climate models using the Earth System Grid Federation data portal (https://esg-dn1.nsc.liu.se/search/cordex/). I then extracted daily minimum and maximum temperature, precipitation, and solar radiation for 1984-2049 under RCP 4.5 and RCP 8.5 from the 18 climate models for the study sites. Lastly, I applied the quantile delta mapping method to correct bias in the daily weather series extracted from the CORDEX climate scenarios. Quantile delta mapping corrects systematic errors in the climate models in relation to historical observations while preserving average changes produced by the climate models (Cannon et al., 2015). Due to inaccessibility of historical climate data at the study sites, I used climate series for 1984-2005 from the NASA POWER (daily solar radiation, and minimum and maximum temperature) and CHIRPS (daily rainfall) datasets as historical observations to conduct bias correction. Bias correction was executed using the "MBC" library of the R package (Cannon, 2018).

Management (particularly planting density) and soil information for each site was identified during the field questionnaire and yield-forecasting algorithm, respectively, described in Chapter 2. Maize was parameterized in the SALUS model to represent a short duration (about 90 days) cultivar, which is commonly planted during long rain seasons in bimodal rain areas in Tanzania (Abate et al., 2017; Nkonya, 1998).

Climate Impact Simulation Experiments

I simulated unfertilized rainfed continuous maize cultivation during the masika rain (long rain) season for 30 years for each of the study sites located in typical bimodal rain regions across Morogoro, Kagera, and Tanga under historical climate and climate change scenarios. A total of 60 sites were included in this study (two sites from Chapter 2 located in southern Morogoro, where the rainfall pattern transitions between unimodal and bimodal, were excluded).

Under climate change, two types of projected climate were used as inputs for the SALUS model: (i) outputs from CORDEX's 18 bias-corrected 18 regional climate models under RCP 4.5 and RCP 8.5, and (ii) delta-method projected climate, where multipliers were applied to climatic variables in the historical record, under the two respective RCPs. RCP 4.5 is a medium greenhouse gas emission path in which radiative forcing would stabilize at 4.5 W/m^2 by 2100, whereas RCP 8.5 is a business-as-usual no-adaptation path in which forcing would reach 8.5 W/m^2 by 2100. Regarding the delta method, I first calculated average percentage differences in the four respective climatic variables (minimum and maximum temperature, solar radiation, and rainfall) between historical climate and the respective 18 climate models, and then adjusted daily weather series for 1990-2019 based on the average percentage difference. This was done for RCP 4.5 and RCP 8.5 for three seasons, long rain (March-May), dry (June-September), and short rain (October-December). Therefore, for each site, a total of 19 climates (18 climates from the CORDEX climate models and the delta-method climate) was used as SALUS model inputs under each of the two RCPs. In SALUS simulations, CO_2 was set at 380 ppm under historical climate, 424 ppm under RCP4.5 and 429 ppm under RCP8.5 (Van Vuuren et al., 2011a; Van Vuuren et al., 2011b).

Statistical Analysis

SALUS Model Accuracy Assessment

I used the Nash-Sutcliffee Efficiency (NSE) index and root mean square deviation (RMSD) between the simulations and observations to quantify the accuracy of the SALUS model in simulating maize yield and plant N uptake. The NSE index ranges from -Inf to 1 and it indicates how well the simulated versus observed values fit the 1:1 line. An NSE index of 1 means the simulation model is 100% accurate, and simulated versus observed values are on the 1:1 line; an

NSE index of 0 means the simulation model is as accurate as using the mean of the observed values; a smaller-than-zero NSE index means the model prediction is worse than using the mean of observations (Moriasi et al., 2007). The RMSD indicates overall error in model prediction and is sensitive to extreme values. The two metrics provide complimentary information for simulation model accuracy evaluation (Willmott, 1981).

Evaluation of Climate Change Uncertainty and Climate Impact Assessment

I used coefficient of variation (CV, ratio between standard deviation and mean) to quantify the impact of climate change uncertainties on three agronomic outcomes, grain yield, SOC, and soil inorganic N. CV was first calculated based on 18 simulations using weather from each of the 18 climate models for each site under each of the two RCP scenarios. I then reported the average and standard deviation of the site-based CV grouped by district and RCP scenario for each of the three agronomic variables. In addition, I compared four climatic variables in long rain seasons (average minimum and maximum temperature, solar radiation and total rainfall) between each of the 18 climate models in 2020-2049 and historical years in 1990-2019 for each site and reported the distribution of the difference grouped by district and RCP scenario.

Changes in 30-year average grain yield under projected climate for 2020-2049 and under historical climate for 1990-2019, changes in SOC over the simulation period 2020-2049, and changes in end-of-30-year soil inorganic N between projected climate and historical climate were of interest in climate impact assessment, in addition to their absolute simulated values. To evaluate the effect of climate change variability on climate impact, I compared the three relevant variables averaging simulated values from 18 climate models versus using one delta-method projected climate. This discrepancy between the average value from 18 simulations and one simulated value based the delta method projected climate indicated inadequacy of using changes

in the mean of climatic variables in climate impact assessment.

Results

SALUS Model Validation

The simulated grain yield matched with the field observations for 62 fields across Tanzania in 2017-2018. The NSE value between the simulated and observed yield was 0.94, indicating the simulated versus observed yields fell closely on the 1:1 line. Overall, the SALUS model captured maize growth under a range of conditions where final observed yield ranged from 0.2 to 5.6 Mg/ha with small error (RMSD value of 0.3 Mg/ha) (Figure 6).

The SALUS model was also able to reproduce maize yield and plant N uptake under various treatments at two sites for 2-4 years in Zheng et al. (2018) and Zheng et al. (2019). The NSE value between the simulated and observed yields under three N rates (0, 50 and 100 kgN/ha) for four years and under 50 kgN/ha with residue addition for two years at two sites was 0.75 (Figure 7a). The RMSD value between yield simulations and observations was 0.6 Mg/ha. For plant N uptake under the three N rates for two years at two sites, the NSE value was 0.88 and RMSD was 11.7 kg/ha (Figure 7b).

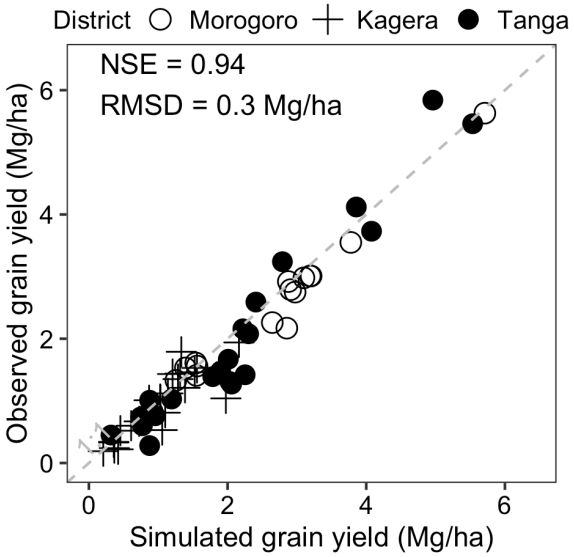

Figure 6 Comparisons between simulated and reported maize grain yield at 62 sites across three districts in Tanzania. (Observations were from Liu and Basso, 2020b)

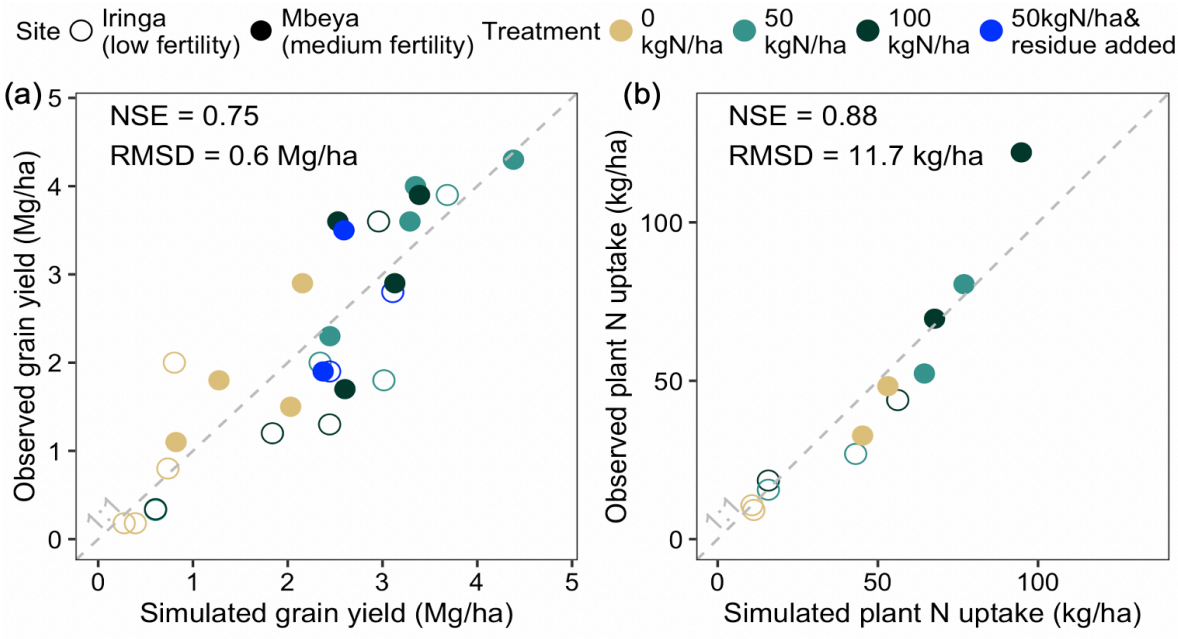

Figure 7 Comparisons between simulated and observed (a) grain yield and (b) plant N uptake under various N-fertilizer and residue incorporation treatments at two locations in Tanzania. (Observations were from Zheng et al., 2019 (ab))

Climate Change Uncertainties

Temperature, rainfall, and solar radiation varied spatially across the study sites (APPENDIX

B Figure 32). Among the four climatic variables, the regional climate models agreed on the projected maximum temperature. The average CV of maximum temperature in the 18 climate models for 60 sites in long rain seasons was about 2-3% under RCP 4.5 and RCP 8.5 (Table 1). Projected total rainfall varied most across the climate models. The average CV of projected total rainfall was about 50% for Morogoro and Tanga under RCP 4.5 and was about 30% for Kagera. The CV was about 30% for Morogoro and about 45% for the other two districts under RCP8.5 (Table 1). For average daily solar radiation and minimum temperature during long rain seasons, CV across the 18 climate models was 2-6% under the two RCPs (Table 1).

Minimum temperature in long rain seasons was projected to increase by the largest extent in Kagera, followed by Tanga and Morogoro. The average daily minimum temperature in 2020-2049 under RCP4.5 was projected to change by -0.05 to 6.7°C with an average value of 1.4°C across the 39 sites and 18 climate models in Kagera. Under RCP 4.5, the average minimum temperature was projected to change by -1.3 to 1.5°C (average of 0.6°C) and -1.4 to 1.8°C (average of 0.2°C) in Tanga and Morogoro, respectively. Under RCP 8.5, the average minimum temperature was projected to change by 0.2 to 7.3°C (average of 1.4°C), -1.3 to 1.8°C (average of 0.8°C), and -1.4 to 2.0°C (average of 0.4°C) in Kagera, Tanga, and Morogoro, respectively (Figure 8a).

Maximum temperature is projected to increase in most cases. The changes in average maximum temperature between 2020-2049 and 2010-2019 ranged from decreasing by 0.6°C to increasing by 4.1°C across the three districts (average change was increasing by about 1.6-1.9°C) under RCP 4.5. Similarly, the changes ranged from decreasing by 0.8°C to increasing by 4.3°C (average change was increasing by about 2°C) under RCP 8.5 (Figure 8b).

Seasonal total rainfall was projected to decrease in most cases. On average, total rainfall was

projected to decrease. It was projected to decrease by 0.5-63.1% (average of 29.0%) in Kagera under RCP 4.5 (Figure 8c). Under RCP 8.5, the percentage change ranged from decreasing by 67.4% to increasing by 2.4% (average of -29.4%). For Morogoro and Tanga, the respective percentage change ranged from decreasing by 67.0% to increasing by 23.3% (average of decreasing by 31.9%), and from decreasing by 64.0% to increasing by 17.2% (average of decreasing by 35.5%) under RCP 4.5. Under RCP 8.5, the percentage changes were from decreasing by 71.1% to increasing by 13.6% (average of decreasing by 32.0%) for Morogoro, and from decreasing by 65.8% to increasing by 8.8% (average of decreasing by 34.9%) for Tanga (Figure 8c).

The percentage change in average daily solar radiation during long rain seasons between the projected climate from the climate models and historical climate was within 10% under RCP 4.5 and was -8% to 12.5% under RCP 8.5 for Kagera. The percentage changes ranged from -10.7% to 19.2% with a respective average of 4.7% and 5.1% under RCP 4.5 and RCP 8.5 for Morogoro. Similarly, the percentage changes were similar under RCP 4.5 and RCP 8.5 in Tanga, ranging from -8% to 12%, with an average value of 1.2% and 1.4%, respectively (Figure 8d).

Table 1 Average and standard deviation of CV (%) of average minimum temperature, average max temperature, total rainfall and daily solar radiation in 2020-2049 across 18 climate models for the simulated 60 sites grouped by the three districts (values in the parentheses are standard deviation)

	Morogoro		Kagera		Tanga	
	RCP 4.5	*RCP 8.5*	*RCP 4.5*	*RCP 8.5*	*RCP 4.5*	*RCP 8.5*
min. temperature	5.12 (2.0)	4.00 (2.6)	1.97 (1.3)	1.83 (1.4)	4.97 (2.0)	3.93 (1.6)
max. temperature	2.04 (1.2)	2.28 (1.7)	3.03 (1.0)	3.09 (1.8)	2.39 (1.5)	1.78 (1.0)
total rainfall	48.18 (16.0)	27.59 (16.8)	29.52 (19.1)	46.89 (22.5)	55.10 (25.7)	41.13 (14.7)
solar radiation	4.51 (2.1)	2.15 (2.2)	5.91 (2.3)	5.38 (2.1)	3.32 (0.9)	3.04 (2.3)

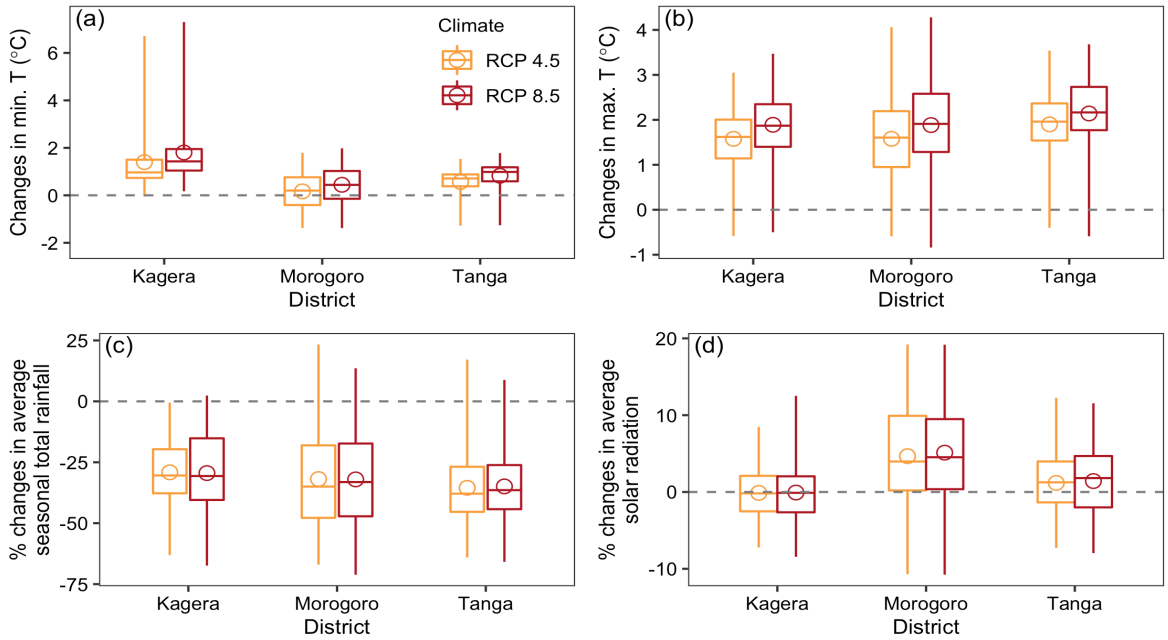

Figure 8 Distribution of changes in average values of climatic variables in masika seasons (March-May) in 1990-2019 across 18 climate models for 60 sites in the three districts under RCP4.5 and RCP8.5: (a) changes in minimum temperature, (b) changes in maximum temperature, (c) percentage changes in seasonal total rainfall, and (d) percentage changes in daily solar radiation.

Uncertainties in Grain Yield, SOC and Soil Inorganic N under Climate Change

The bias-corrected weather input from the 18 climate models led to substantial variations in simulated soil inorganic N (CV ranging from 5% to 283%), sizable variations in grain yield (CV: 3-21%) and negligible variation in SOC (CV within 4%) (Figure 9).

Across the study sites, the average CV of simulated 30-year average yield in 2020-2049 was 8-9% for Morogoro and Kagera under RCP 4.5 and RCP 8.5, and was about 16% for Tanga under the two RCPs. In addition to uncertainties attributed to projected climate, spatial variability accounted for additional uncertainties. The standard deviation of grain yield simulation uncertainty was about 4% (Figure 9a).

The effect of variability among climate models was the least on SOC simulations. Across the study sites under RCP 4.5, the average and standard deviation of CV of simulated end-of-30-years SOC was 1.1% and 0.5%, respectively for Morogoro, 0.9% and 0.3% for Kagera, and 1.5%

and 0.9% for Tanga. Under RCP 8.5, the CV slightly increased, with average values of 1.0%, 1.0% and 1.6% for Morogoro, Kagera, and Tanga, respectively (Figure 9b).

In general, the uncertainty in soil inorganic N simulation was the largest due to the uncertainties in heavy rain events under the projected climate. On average, CV of simulated soil N at the end of 30-year simulations was 89-117% for the three districts under RCP 4.5, and was smaller under RCP 8.5, ranging from 47% to 74%. In addition, there was a large spatial variability in soil N simulations. The standard deviation of CV of simulated soil N across the sites ranged from 52% to 97% under RCP 4.5 and was 31-66% under RCP 8.5 (Figure 9c).

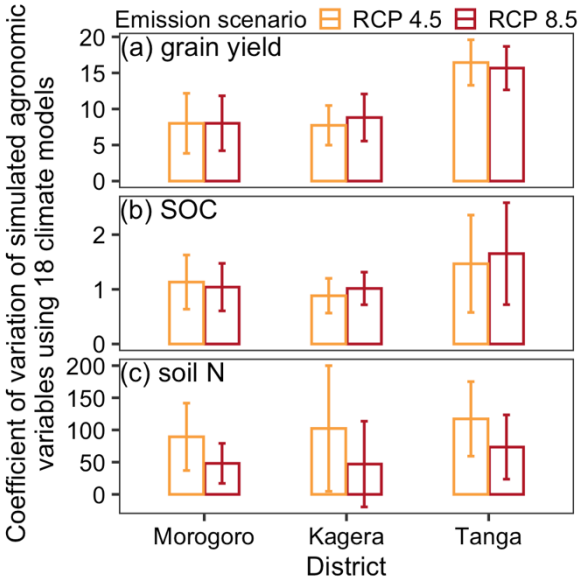

Figure 9 CV of simulated agronomic variations when using output of 18 climate models as SALUS input across the study sites in three districts under RCP 4.5 and RCP 8.5: (a) grain yield, (b) SOC and (c) soil inorganic N (note the ranges of y axis differ for the three panels).

Discrepancy in Climate Impact Assessment

Across each study site under RCP 4.5 and RCP 8.5, the predicted changes in 30-year average yield due to climate change varied substantially across the 18 climate models (APPENDIX B Figure 33). When averaging the change simulated using the 18 climate models as SALUS input for each site, Morogoro was predicted to have decreased yield (by 1.0-19.3%) for nine sites, and

to have increased yield (by up to 7%) for the remaining seven sites under the RCP 4.5 scenario. Similar values were found for RCP 8.5 for study sites in Morogoro. Under RCP 4.5, yield in Kagera was predicted to decrease by 0.6-40.8% for 23 sites, and one site was predicted to increase insignificantly (by 0.05%). Under RCP 8.5, yield was predicted to decrease more, by 2.0-44.1% for 21 sites, and to increase by up to 10%. For Tanga, on average for the 18 climate models, yield was predicted to decrease by about 15-48% under the two RCP scenarios (Figure 10).

When using one delta-method projected climate, the simulated percentage change tended to be smaller than the averaged simulated change using the 18 climate models for the study sites. For Morogoro, the simulated percentage change was about 6% smaller than the average simulated changes using 18 climate models under the two RCPs. For Tanga, such discrepancy was 17-18% under RCP 4.5 and RCP 8.5 scenarios. The simulated percentage change in yield, nonetheless, was close to the average simulated percentage changes using 18 climate models for study sites in Kagera district under both RCP scenarios (Figure 10).

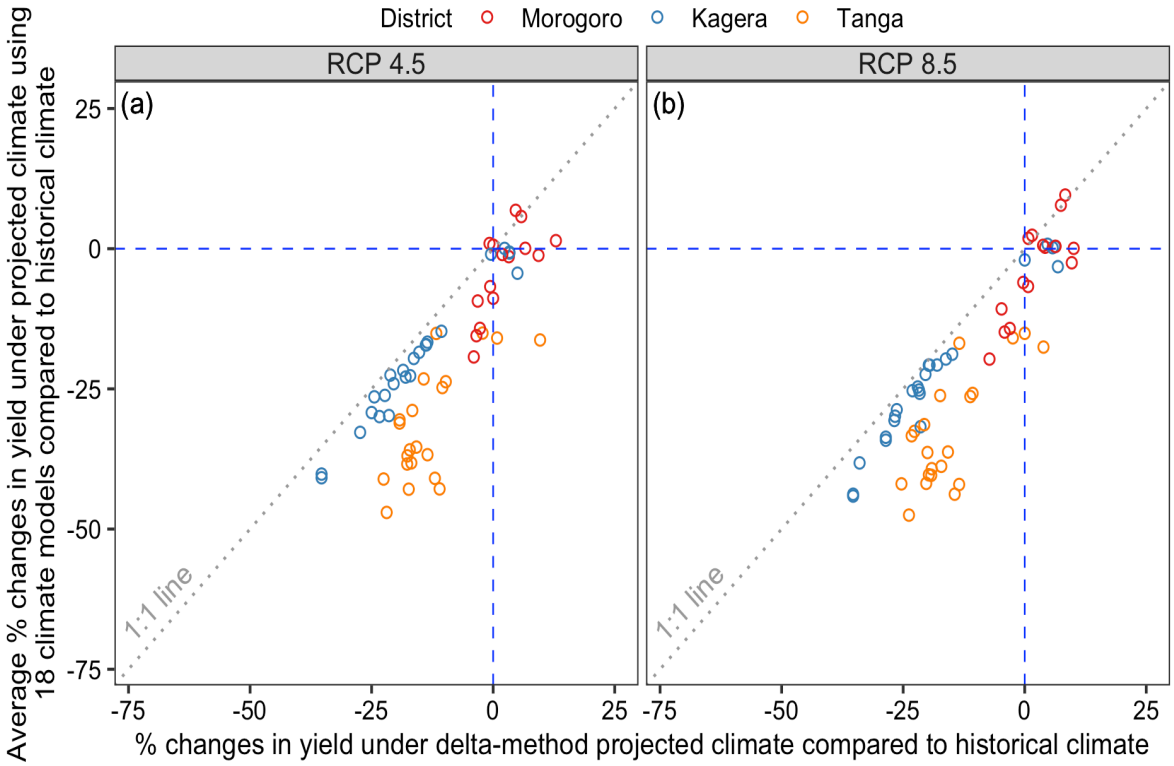

Figure 10 Comparisons of average simulated percentage changes in yield due to climate change using 18 climate models versus using one delta-method projected climate for (a) RCP 4.5 and (b) RCP 8.5

The average SOC change using the 18 climate models was close to that using the delta-method climate for both RCPs for most sites, with the exception of three sites in Tanga where the average SOC from the models decreased slightly more than that of the delta-method climate. Averaging SOC changes from the 18 climate models, SOC was predicted to decrease by 13.1-31.4% (average of 23.2%) across the study sites under RCP 4.5 and by 13.3-31.8% (average of 23.6%) under RCP 8.5 (Figure 11, APPENDIX B Figure 34).

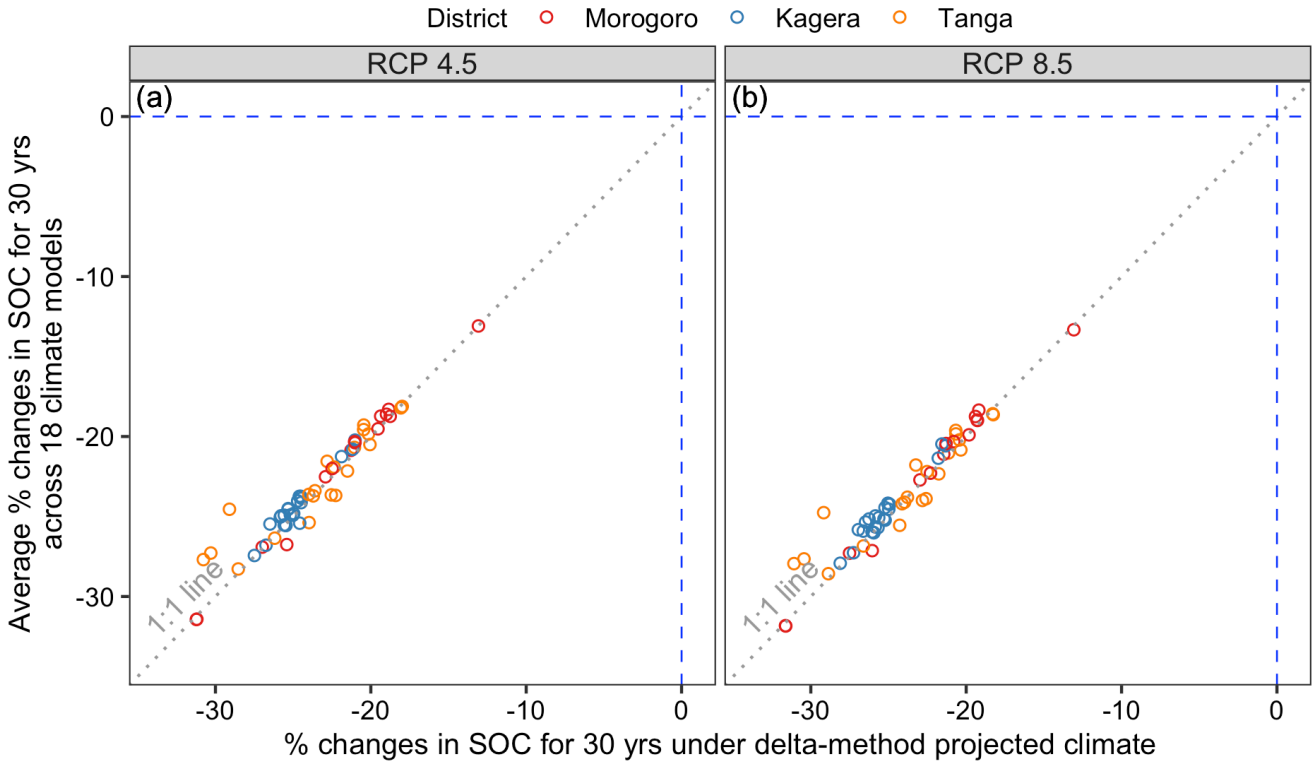

Figure 11 Comparisons of average simulated SOC changes over 30 years using 18 climate models versus using one delta-method projected climate

The difference between simulated soil inorganic N under projected climate and under historical climate was highly dependent on the weather inputs of the SALUS model (APPENDIX B Figure 35). Averaging the simulated changes in soil inorganic N using the 18 climate models, soil inorganic N was predicted to increase by up to 6 kg/ha for 14 sites in Morogoro under RCP 4.5, and by up to 2 kg/ha under RCP 8.5 for 13 sites. One site was predicted to have decreased soil inorganic N under RCP 4.5 and soil N would be lower for two sites under RCP 8.5 for Morogoro (both by less than 0.5 kg/ha) (Figure 12). For Kagera, all sites were projected to see increased soil inorganic N, by up to 13 kg/ha, under climate change when averaging the simulated changes across the 18 models under both RCP scenarios. Similar findings apply to study sites in Tanga, with up to 14 kg/ha increase in soil N under projected climate (Figure 12). When using one delta-method projected climate, the simulated change in soil inorganic N was

close to the averaged simulated change using the 18 climate models (within 5 kg/ha difference) for most sites across the three districts, particularly in Morogoro and Kagera. Such difference was more than 5 kgN/ha (up to 14 kg/ha) for 14 sites in Tanga under the RCP 4.5 scenario, and for 5 sites under the RCP 8.5 scenario (Figure 12).

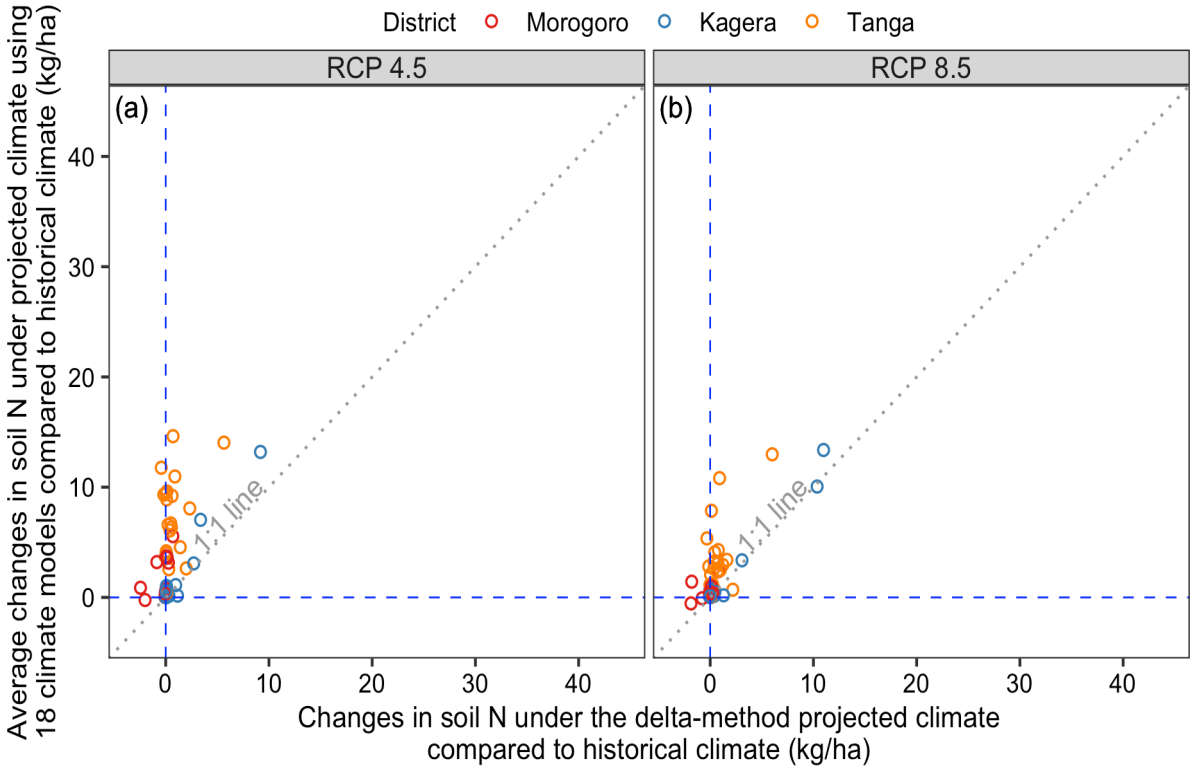

Figure 12 Comparisons of average simulated changes in soil inorganic N under climate change compared to historical climate when using 18 climate models versus using one delta-method projected climate

Discussion

Results from analyzing changes in climatic variables in 2020-2049 compared to 1990-2019 aligned with findings in the climate change literature that future temperatures will continue to warm, but much uncertainty remains in precipitation projection (Pfahl et al., 2017; Sillmann et al., 2017). The variations in the climate projected by 18 regional climate models led to varied projections of climate impact for smallholder maize fields, using regional climate model output

to drive the SALUS crop model. The variability of simulated yield and soil inorganic N was sizable among the climate models, but the variability in SOC was low (Figure 9). Thus, this study provided insights on the need for multimodel ensemble of regional climate models as crop model inputs for climate change impact analysis.

Several studies have evaluated climate impact on maize yield in Tanzania. In general, maize yield has been projected to decrease due to shorter growing seasons caused by increased temperatures, and/or drought stress caused by decreased rainfall (Rowhani et al., 2011). The findings in this study support the prediction of decreased maize yield under climate change. But the results in this paper also suggest that depending on the regional climate model output used in a crop model, maize yield could be projected to increase, particularly for sites in Morogoro (Figure 10, APPENDIX B Figure 33). Others have also reported increased maize yield under the projected climate in Morogoro (by less than 10%) and Kagera (by less than 2%) (Arndt et al., 2012; Luhunga et al., 2017; Luhunga, 2017). Moore et al. (2012) attributed such increase in yield under climate change to more favorable conditions, and in addition, Folberth et al. (2016) showed that soils used in crop models may influence the direction of projected changes.

This study contributed to the inclusion of soil-management-crop feedback in climate impact assessment. Compared to historical climate, soil inorganic N was projected to increase under the projected climate, but the increased inorganic N — which resulted from higher mineralization and less plant uptake due to smaller plant size under climate change — does not linearly translate to higher grain yield. This finding was in line with one of the few climate impact studies on soil inorganic N (Basso et al., 2018b).

Incrementally adjusting crop models' climatic inputs — temperature, precipitation, and CO_2 — is one common practice to evaluate climate change impact on cropping systems. This practice

does not account for changes in climate extremes. Here, the results — comparing averaged climate impact from 18 climate models on various aspects of maize production systems to climate impact using a delta-method projected climate that only considered changes in the mean of climatic variables — indicated that yield and soil inorganic N were sensitive to climate extremes but soil carbon was not (Figure 10-12).

Conclusions

Evaluating uncertainties in projected climate and in climate change impact on agriculture is important for planning agronomic strategies able to allow crops to adapt to a new climate regime. This paper addressed uncertainties in climate impact on yield, SOC, and soil N of maize-based systems in Tanzania resulting from variability in the projected future climate. The results showed that maize yield and SOC would most likely decrease, and soil N would increase. The results also implied the importance of considering changes in climate extremes in climate impact assessment.

Acknowledgements

This work was made possible through support provided by the US Agency for International Development (USAID) under the terms of Cooperative Agreement No. 7200AA18LE00003, with additional funding from US Department of Agriculture National Institute of Food and Agriculture (award no. 2015-68007-23133), US National Science Foundation's Kellogg Biological Station Long Term Ecological Research Site (NSF grant no. DEB 1027253), and Michigan State University AgBioResearch and USDA/NIFA HATCH grant No. MCL02368.

CHAPTER 4: SUSTAINABLE YAM PRODUCTION IN GHANA: A MODELING PERSPECTIVE ON THE RESPONSE OF YAM YIELD AND SOIL ORGANIC CARBON TO AGRONOMIC MANAGEMENT

A version of this chapter has been submitted to a peer-review journal

Abstract

Yam, a major food crop for West Africa, has not been managed to reach its potential productivity. The current practice of planting yam continuously for years after clear-cutting a field is not sustainable and has led to deforestation and nutrient depletion. By examining the effect of improved management on yam cultivation in Ghana, this study aimed to solve the tradeoff between improving yam yield and sustaining soil organic carbon (SOC).

We used a calibrated and validated process-based crop model, Systems Approach to Land Use Sustainability, to assess the impact of four management treatments: continuous unfertilized rainfed yam (control), pigeonpea-yam rotation, yam with 3 Mg/ha pigeonpea residue incorporated and yam with 23-23 N-P2O5 kg/ha fertilizer added. We modeled 10 years of yam yield and SOC across cropland in Ghana with varying levels of soil carbon, rainfall amount, and precipitation pattern. On average, simulated yam tuber yield was the highest with a pigeonpea residue incorporation treatment (4.1-11.9 Mg/ha, average of 7.5 Mg/ha). The rotation (average yield of 6.4 Mg/ha) and fertilizer (average of 7.0 Mg/ha) treatments produced comparable increases in yam yield over the control treatment (1.9-9.2 Mg/ha, average of 4.9 Mg/ha). The low yam yield of the control treatment was mostly attributed to nutrient deficiency (nitrogen and phosphorus). Drought also limited yam growth, particularly in northern Ghana. The three improved management treatments increased soil nutrient availability and thus improved yield.

SOC declined under all four tested treatments over the simulated 10 years, but declined least with residue incorporation (average rate -0.3 Mg/ha/year), followed by fertilizer addition (-0.43 Mg/ha/year), rotation (-0.42 Mg/ha/year), and the control (-0.51 Mg/ha/year) management. Our work provides a benchmark for yam yield response to alternative management across Ghana, and highlights pigeonpea's contribution to sustainable intensification of yam. Further research is needed to untangle the interacting effects of land use and agronomic management on SOC.

Introduction

White yam (*Dioscorea rotundata*) is an extensively cultivated and consumed species in Ghana and other West African countries (Kayode et al., 2017; Olatoye and Arueya, 2019; Raymundo et al., 2014). The conventional shifting agriculture practice, in which yam is cultivated continuously for 10 years after clear-cutting a field, not only leads to yield decline over time but also exacerbates deforestation (Acheampong et al., 2019; Maliki et al., 2012). To sustainably intensify yam cultivation, agronomic practices need to improve yam yield while also enhance soil fertility.

Yam has yielded below its potential, despite its critical role in food security in Ghana. Average fresh yam tuber yield in 2016 was 17.42 Mg/ha (equivalent to dry tuber yield of 6.1 Mg/ha, given 65% moisture content) in Ghana, despite yield potential of 52 Mg/ha (18.2 Mg/ha dry yield) (MoFA, 2017). Previous research has focused on testing yam tuber yield response to both synthetic and organic fertilizer addition. Several field experiments in West Africa have reported tuber yield under various levels of NPK fertilizer input, as well as yield with different forms of fertilizer, including poultry manure, crop residue, synthetic fertilizer, and combinations of organic and inorganic fertilizers (Agbede et al., 2013; Law-Ogbomo and Egharevba, 2009). In

recent years, pigeonpea, a legume crop, has been introduced to yam-based smallholder fields in Ghana. Few studies have been conducted to determine the effect of pigeonpea on yam yield (Acheampong et al., 2019; Adjei-Nsiah, 2012). While existing research has shown the potential of integrating pigeonpea into cropping systems in Ghana as well as the positive effects of fertilizer on yam yield, it is not well-understood how much yam growth is constrained by soil nitrogen (N) or phosphorous (P), given the limited spatial coverage of existing field studies of Ghana (Carsky et al., 2010; Frossard et al., 2017).

The role of soil organic carbon (SOC), another key component of tropical agricultural systems, has not been explored for yam production in Ghana. Several studies have raised concerns of nutrient depletion and carbon loss under traditional continuous yam production (Abdoulaye et al., 2014; Anikwe, 2010; Carsky et al., 2010). Field studies across various West African countries have shown declines in SOC under continuous cultivation of sorghum, pearl millet, and millet-groundnut intercropping (Bationo et al., 2007; Ouédraogo et al., 2007). Long-term field experiments in Niger have suggested SOC can be maintained for 25 years and even increased under millet-cowpea rotation with the addition of crop residue and synthetic fertilizer (Nakamura et al., 2011). Indeed, improving crop yield, increasing the return of nutrients to soil, and applying soil amendments (such as synthetic fertilizer, crop residue and manure) is critical to maintaining SOC in West Africa (Bationo et al., 2007).

Crop modeling is a valuable tool for efficiently testing the impact of agronomic management on various aspects of cropping systems, including yield and SOC, over large areas and long time spans (Liu and Basso, 2020a; Whitbread et al., 2010). A few crop models—including the CROPSYST and EPIC models—have been adapted to simulate yams in Guadeloupe and Benin (Marcos et al., 2011; Raymundo et al., 2014; Srivastava and Gaiser, 2010). No crop modeling

studies have been conducted to assess the impact of agronomic management on yam yield or SOC in Ghana.

In this study, we investigated the effect of agronomic management on yam yield and SOC across Ghana's diverse landscape, with varying levels of soil fertility and precipitation. The objectives of this paper were: (1) to evaluate the impact of commonly practiced continuous rainfed unfertilized yam cultivation as well as three improved management treatments on yam yield and SOC across Ghana, (2) to identify abiotic yield-limiting factors (water, P and N) for yam across Ghana, and (3) to quantify the contributions of the improved management treatments in reducing yield-limiting stress levels.

Materials and Methods

Study Sites

Ghana is located in the West Africa along the Gulf of Guinea and Atlantic Ocean. The country is made up of tropical savanna with a dry winter climate and six distinct agroecological zones: rainforest, deciduous forest, transitional zone, coastal savanna, Guinea savanna and Sudan savanna (Figure 13a, FAO, 2005). Though studies define agroecological zones slightly differently (e.g. Abbam et al., 2018; Adams et al., 2019; Amekudzi et al., 2015; Rhebergen et al., 2016), Ghana can be divided into four zones: forest in the south, two savanna zones on the coast and in the north of the country, respectively, and a transitional zone in between the forest and the northern savanna zones (Figure 13b). From the southwest to northeast of the country, there is a decrease in annual total precipitation (Figure 13b, Funk et al., 2015). A unimodal rain season occurs between June to September in the northern savanna zone, whereas bimodal rain seasons, April-July and September-November, occur in the two zones in the south and the transitional

zone (Bellon et al., 2020; MoFA, 2017). The predominant soils in Ghana are Ferralsols, Acrisols, Luvisols and Plinthosols. Soil texture is most likely to be sandy loam or loam (MoFA, 2017; Tetteh et al., 2016). Soil organic carbon content varies across the six agroecological zones, ranging from 0.6% in Sudan savanna to 2.7% in the forest zones (Adjei-Gyapong and Asiamah, 2002; Owusu et al., 2020).

Yam is extensively cultivated in Ghana, particularly in three administrative regions in the forest-savanna transitional zone (Brong Ahafo, Bono East, and Ahafo) and in three regions in the savanna zone of northern Ghana (Savannah, Northern and North East regions) (Figure 13a, MoFA, 2017).

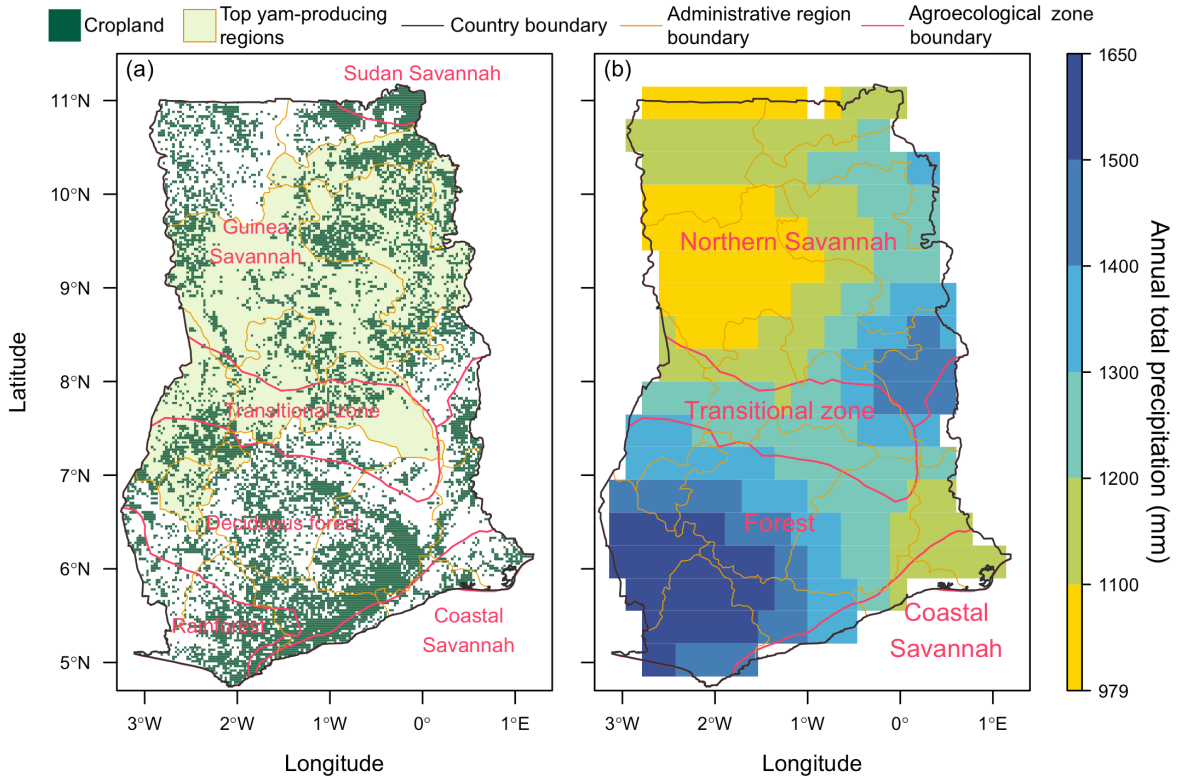

Figure 13 Study sites in Ghana; (a) distribution of cropland overlaid with administrative regions and six agroecological zones (Fritz et al., 2015), and (b) average annual precipitation (Funk et al., 2015).

We classified as cropland (Figure 13a) any area with more than a 52% probability of being cropland based on the global cropland product from Fritz et al. (2015). We chose 52% as the threshold because the resulting map matched the amount of estimated cultivated land area (63,419.3 km^2) in Ghana (MoFA, 2017). The simulation unit was the soil grid at 1 km resolution. A total of 60,554 simulation units were included in the study, thus equating to 60,554 km^2.

SALUS-Yam Model Parameterization and Validation

We tested the SALUS-Yam model for simulating yam phenology and biomass response to N and P fertilizer input against observations in field experiments in Ghana. We compared simulated yam phenology (emergence, senescence, and maturity) and biomass (including tuber, vine, and leaf biomass) to observations across six treatments in two different ecological settings (the towns of Ejura and Fumesua) for two continuous years (2018-2019). These treatments combined two pigeonpea residue application treatments with three fertilizer application rates. The two residue application treatments were yam without residue applied, and yam with about 2 Mg/ha of pigeonpea residue added. The three N-P_2O_5-K_2O fertilizer application rates were 0-0-0, 23-23-30 and 45-45-60 kg/ha. The Pona yam cultivar was planted in both locations. Ejura is located in the forest-savanna transition zone in central Ghana, with relatively low soil organic matter (0.6% in the top 20cm soil layer) and a drier climate (an average of 1250mm of total rainfall in 2018-2019). Fumesua is located to the south of Ejura in the forest zone, with higher soil organic matter (0.8% in the top 20cm) and a wetter climate (an average 1553mm of precipitation in 2018-2019). A total of 24 field observations (2 years x 2 locations x 2 residue treatments x 3 fertilizer treatments) were available for SALUS-Yam model testing.

To parameterize SALUS-Yam, we first randomly selected half of the field observations, and then compared the simulated phenology and biomass (both tuber yield and aboveground

biomass) to the remaining half of field observations. The yam parameters for the SALUS simple mode were based on reported values in the literature and the calibration processes (Table 2). Weather information used in SALUS model testing was obtained from weather stations at the two locations. Soil and management information were derived from the field experiments. We also applied a yield reduction factor to the simulated tuber and aboveground biomass based on observed weed levels (APPENDIX C Figure 36).

Table 2 Critical crop parameters and the parameter value used in the SALUS-Yam model

Parameter	Descriptions (unit)	Value	References
TbaseDev	Base temperature (°C)	15	Srivastava and Gaiser (2010)
ToptDev	Optimal temperature (°C)	30	Srivastava and Gaiser (2010)
TTtoGerm*	Degree days required for germination	400	
TTtoMatr*	Degree days required for maturity	2350	
RelTT_P1*	Relative thermal time near emergence	0.15	
RelTT_P2*	Relative thermal time near senescence	0.8	
RelLAI_P1*	Relative LAI near emergence	0.05	
RelLAI_P2*	Relative LAI near senescence	0.95	
RUEmax	Maximum radiation use efficiency (g/MJ)	2	Diby et al. (2011)
LAImax	Maximum leaf area index	6	Law-Ogbomo and Remison (2008)
GrnN_Mt	Tuber N concentration at maturity	0.01	Frossard et al. (2017)
PlntN_Em*	Plant N concentration at emergence	0.025	
PlntN_Hf	Plant N concentration at mid-season	0.02	Law-Ogbomo and Remison (2009)
PlntN_Mt	Plant N concentration at maturity	0.01	Law-Ogbomo and Remison (2009)
GrnP_Mt	Tuber P concentration at maturity	0.0008	Frossard et al. (2017)
PlntP_Em*	Plant P concentration at emergence	0.0022	
PlntP_Hf	Plant P concentration at mid-season	0.0018	Law-Ogbomo and Remison (2009)
PlntP_Mt	Plant P concentration at maturity	0.001	Law-Ogbomo and Remison (2009)

*Values were based on field experiment observations.

Simulation Experiments and Inputs

We simulated yam cultivation for 10 years (2010-2019) under various agronomic practices on each simulation unit (soil pixels at 1km resolution) across cropland in Ghana using the validated yam crop parameters and gridded NASA POWER weather and Africa Soil Information Service (AfSIS) soil datasets.

Agronomic practices in this study included the traditional practice of continuous 10-year unfertilized and rainfed yam cultivation, and three improved practices, (i) continuous yam

cultivation with 3 Mg/ha of pigeonpea residue added (ii) pigeonpea-yam rotation, and (iii) continuous yam cultivation with 23 kg/ha N-P2O5 added (23 kg/ha N-P2O5 was half of the government recommended yam fertilization rate, and equivalent to 23 kgN/ha and 10 kgP/ha) (Table 3). For all tested agronomic practices, yam was planted at a density of 1 plant/m^2. Planting time was determined by the beginning of the rainy season. To ensure adequate plant available water, the beginning of the rainy season was defined as when it rained for five consecutive days and cumulative rainfall reached 25 mm (typically July for the northern savanna zone, and late April to early May for the other zones). For the simulations of continuous yam cultivation with pigeonpea residue addition, pigeonpea leaf N concentration was set at 2.5%, and leaf P concentration at 0.15% (Kesh et al., 2017; Phiri et al., 2010). A medium-duration annual pigeonpea (6-month duration) was used in the 2010-2019 pigeonpea-yam rotation. Pigeonpea crop parameters were drawn from our previous study in Malawi (Liu and Basso, 2017a). Yam residue (vines and leaves) was left in the field in improved practice simulations, whereas no yam residue was left in the field after harvest under control management.

Table 3 Descriptions of the tested management treatments in this study (all treatments were rainfed and were simulated continuously for 10 years)

Treatment	Descriptions
Control (C)	Unfertilized continuous yam cultivation, a common practice among local farmers
Residue (Res)	Continuous yam with 2Mg/ha pigeonpea leaf residue incorporated
Rotation (Rot)	Pigeonpea and yam rotation
Fertilized (F)	23 kg/ha N-P_2O_5 fertilizer added to continuous yam (equivalent to 23 kgN/ha and 10 kgP/ha)

Weather data for 2010-2019 was drawn from the gridded POWER agroclimatology dataset. The POWER dataset provides daily minimum and maximum temperature, precipitation, and solar radiation data from across the globe at 0.5° spatial resolution (data was downloaded from https://power.larc.nasa.gov). We did not preprocess POWER weather series data to remove

drizzling precipitation (<1mm/day) or correct bias for climatic variables for two reasons. Annual total precipitation in POWER weather dataset in general agreed with the total precipitation recorded by weather stations ($r^2 = 0.8$) (APPENDIX C Figure 37a). Adjusting POWER data had minimal impact on yield simulations (APPENDIX C Figure 37b).

We extracted soil organic carbon, bulk density, and silt and clay content data at various depths (5, 15, 30, 60, and 100 cm) from a gridded soil database with 1km spatial resolution developed by the AfSIS (Hengl et al., 2014). We used the database's medium estimate for the soil parameters. We resampled the AfSIS's 250m-resolution top layer (top 30cm) extractable P dataset to 1km-resolution to initialize P for each soil pixel simulation unit (Hengl et al., 2017). In our simulation experiment, the intermediate SOC pool was initialized using procedures in Basso et al. (2011).

Statistical Analysis

To evaluate SALUS-Yam model accuracy for simulating yam growth in Ghana, we calculated RMSD and mean absolute percentage error (MAPE) between the simulations and the observations for calibration and validation datasets.

We compared the 10-year average simulated yam tuber yield under each of the tested management treatments across the 60,554 simulation units in Ghana. We then compared average and standard deviation of the simulated average yam tuber yield grouped by agroecological zone and treatment.

We used Kruskal-Wallis one-way analysis of variance to identify which abiotic factors (water, N or P deficiency) constrained rainfed unfertilized yam growth. The deficiency stress level was calculated based on Equation (3) and Equation (4). We reported the stress factor with the computed stress level statistically higher than the other two at $P = 0.05$. We reported two

stress factors if the two stress levels were statistically higher than the third factor but were indistinguishable from each other at P = 0.05. In case of no statistical differences among the three stress levels, we reported all three stress factors. We applied the nonparametric Kruskal-Wallis test to three stress levels for each simulation unit due to the skewed distribution of the stress levels (APPENDIX C Figure 38). In addition, for each factor, we quantified the level of stress reduction resulting from the improved management treatments compared to the control treatment. To do this, we computed the relative difference in the average simulated 10-year stress factor for each constraining factor and each simulation unit.

$Stress\ Factor_i = Supply/Demand$…… Equation. (3),

$Stress\ Level = \sum_{i=1}^{i=n} 1 - Stress\ Factor$…..…….....… Equation (4),

where the *Supply* and *Demand* are simulated daily supply and demand of resource (water, N and P), respectively, *i* denotes *i*th day, and *n* denotes total number of days in a growing season.

To compare SOC response to management treatment, we first reported changes in SOC in 26 cm of soil over the course of the 10-year simulation for each simulation unit, and then reported average and standard deviation of SOC change by agroecological zone and treatment.

Results

SALUS Model Calibration and Validation

For both field years and locations, yam emerged 6 weeks after planting, and senescence occurred 28 weeks after planting. Yam was harvested in December. In the calibration dataset, the simulated emergence came 44 days (6 weeks) after planting and senescence was 26-29 weeks

after planting. Simulated maturity occurred in early to mid-December. Similar phenological stages were simulated for the validation dataset (Table 4).

Table 4 Comparison between reported and SALUS-Yam simulated yam phenology

Phenology	Observed	Simulated	
		Calibration	Validation
Emergence	42 DAP[1]	44 DAP	44 DAP
Senescence	28 WAP[2]	26-29 WAP	26 WAP
Harvest/maturity[3]	December	Early to mid-December	Early to mid-December

[1]DAP: days after planting
[2]WAP: weeks after planting
[3]Simulated maturity date and reported harvest time

The simulated yam tuber yield closely matched reported yield across treatments: with and without pigeonpea residue addition and with the three different rates of N and P fertilizer inputs. Using the calibration dataset, RMSD between the simulated and the observed tuber yield was 0.8 Mg/ha and MAPE was 11.1% (Figure 14a). With the validation dataset, RMSD between simulated and observed tuber yield was 0.9 Mg/ha and MAPE was 15.0% (Figure 14b). For the aboveground biomass simulation, RMSD between simulations and observations was 0.1 Mg /ha for both calibration and validation datasets, and the MAPE values were 12.1% and 11.3%, respectively (Figure 15).

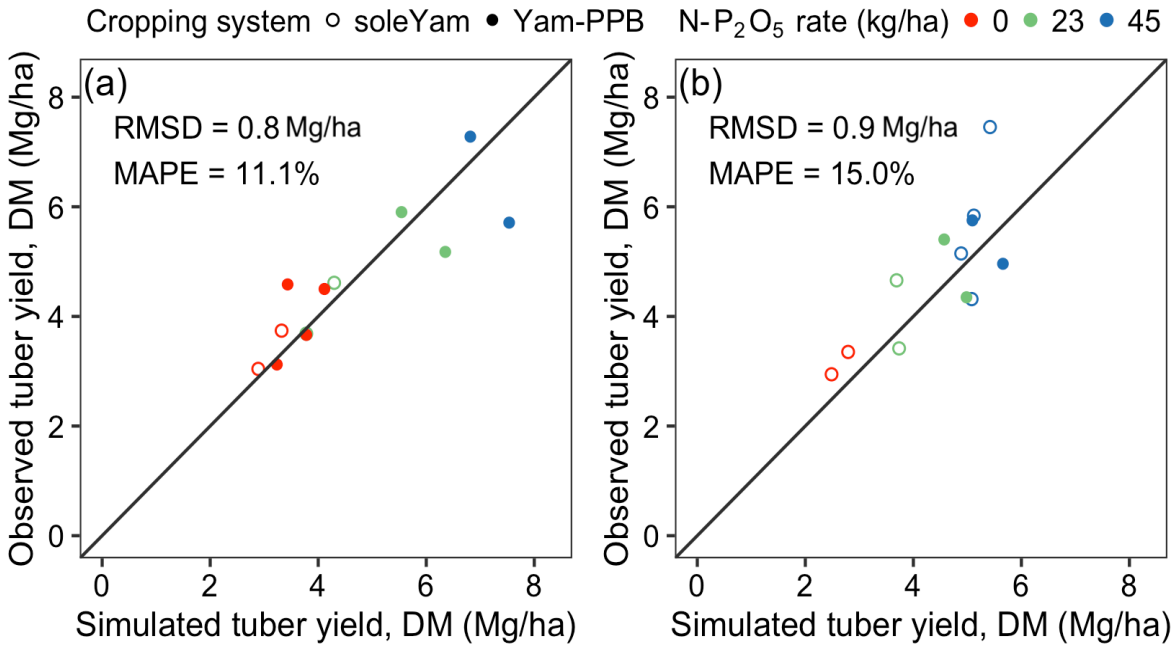

Figure 14 Comparisons between simulated tuber dry matter (DM) yield and observed yield using (a) calibration dataset and (b) validation dataset (abbreviation for cropping systems: soleYam means the sole yam; Yam-PPB means yam with pigeonpea at the border).

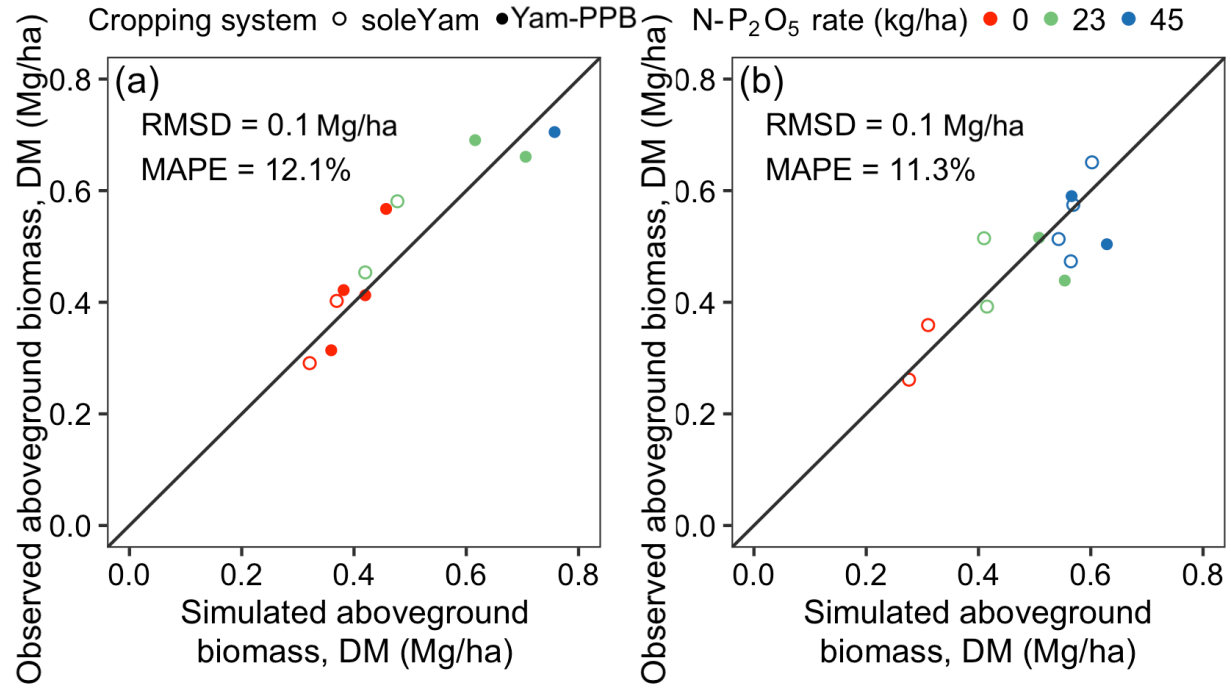

Figure 15 Comparisons between simulated aboveground (leaf and vine) dry matter biomass and observed aboveground biomass using (a) calibration dataset and (b) validation dataset (abbreviation for cropping systems: soleYam means the sole yam; Yam-PPB means yam with pigeonpea at the border).

Yam Tuber Yield Response to Management Treatment

When yam was managed under the control (rainfed and unfertilized) condition, simulated 10-year average tuber yield across Ghana ranged from 1.9 to 9.2 Mg/ha, averaging of 4.9 Mg/ha, with the highest yield in the forest agroecological zone (an average of 6.0 Mg/ha and standard deviation of 0.6 Mg/ha), followed by the coastal savanna zone (average of 5.0 Mg/ha and standard deviation of 0.5 Mg/ha), transitional zone (average of 5.0 Mg/ha and standard deviation of 1.0 Mg/ha), and northern savanna zone (average of 3.6 Mg/ha and standard deviation of 0.8 Mg/ha) (Figure 16a-b).

Yam yield improved when pigeonpea or synthetic fertilizer was added to the yam-based cropping system. Among the three tested improved management treatments, yam yield was enhanced by the largest amount under the pigeonpea residue incorporation treatment. With pigeonpea residue incorporation, yam yield increased to 4.1-11.9 Mg/ha (with an average of 7.5 Mg/ha), which was 21.5%-130.3% more than yield under the control treatment (Figure 16c). Average yield with pigeonpea residue incorporation in the forest zone was 8.8 Mg/ha, 7.8 Mg/ha in coastal savanna, 8.0 Mg/ha in transitional zone, and 5.9 Mg/ha in northern savanna (Figure 16a). Under pigeonpea-yam rotation treatment, yam yield was 3.0-11.1 Mg/ha across Ghana with an average of 6.4 Mg/ha (Figure 16d). Average yam yield under the rotation treatment in forest zone was 7.6 Mg/ha, about 6.5 Mg/ha in the transitional and coastal savanna zones, and the lowest (4.9 Mg/ha) for the northern savanna zone (Figure 16a). Adding 23 kg/ha of N-P2O5 fertilizer generated slightly higher tuber yield than the rotation treatment. The 10-year average yield with fertilizer addition ranged from 3.7 to 11.1 Mg/ha, averaging at 7.0 Mg/ha (Figure 16e). A similar spatial pattern of yield response was observed. Yam yield decreased from the southwest forest zone to the northern savanna zone, with an average yield of 8.1 Mg/ha in the

southwest forest zone, about 7.0 Mg/ha for the coastal savanna and transitional zones, and 5.5 Mg/ha in the northern savanna zone (Figure 16a).

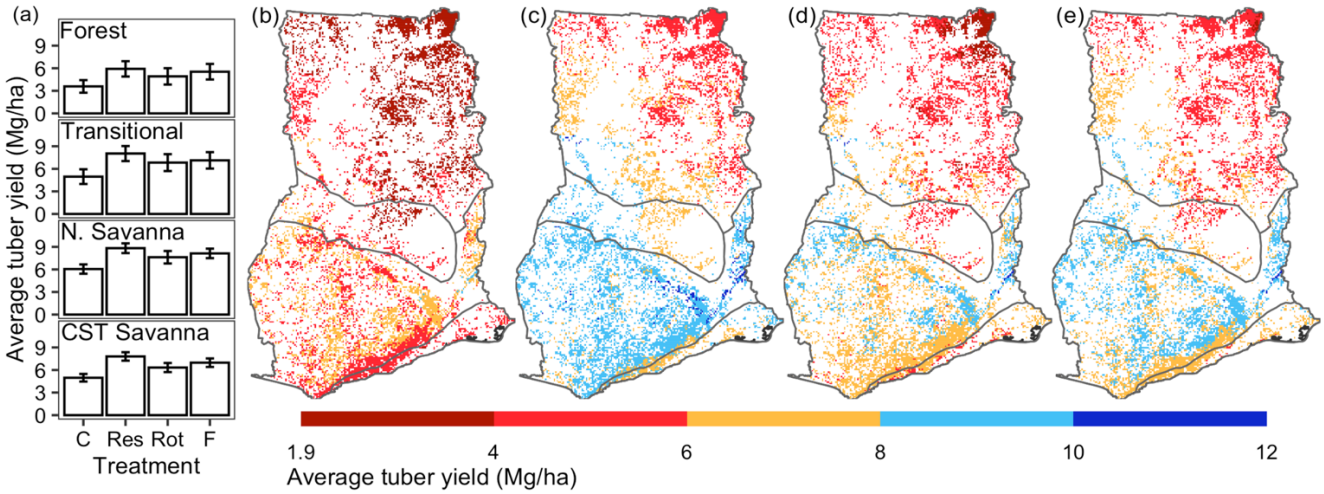

Figure 16 Tuber yield response to management treatments across Ghana. (a) 10-year average tuber yield in the four agroecological zones under the four tested management treatments, and (b-e) 10-year average tuber yield under control, pigeonpea residue incorporation, pigeonpea-yam rotation, and fertilizer addition treatment, respectively. Abbreviations in (a): N. Savanna – northern savanna, CST savanna – coastal savanna, C – control, Res – residue, Rot – rotation, and F – fertilized. Error bars in (a) represent standard deviation.

Abiotic Limiting Factors for Yam Tuber Yield

Yam tuber yield under rainfed unfertilized cultivation was mostly constrained by lack of N, the dominant yield-limiting factor across each of the four agroecological zones. N deficiency alone limited yam yield for 92%, 82%, 77%, and 69% land area of the coastal savanna, transitional, forest and northern savanna zones, respectively (Figure 17a-b). N and P co-limitation counted for an additional 8%, 17%, 23%, and 8% of the land area of the respective zones. Yam yield was mostly constrained by water deficiency for 11% of the land area in northern savanna, located in the west of the northern savanna zone. An additional 12% of land area in the northern savanna was co-limited by N and water. Yam tuber yield was constrained by all three stress factors for an insignificant fraction of land area (0.4%) in the northern savanna,

where no statistical differences were found for water, N, and P stress levels (Figure 17a-b).

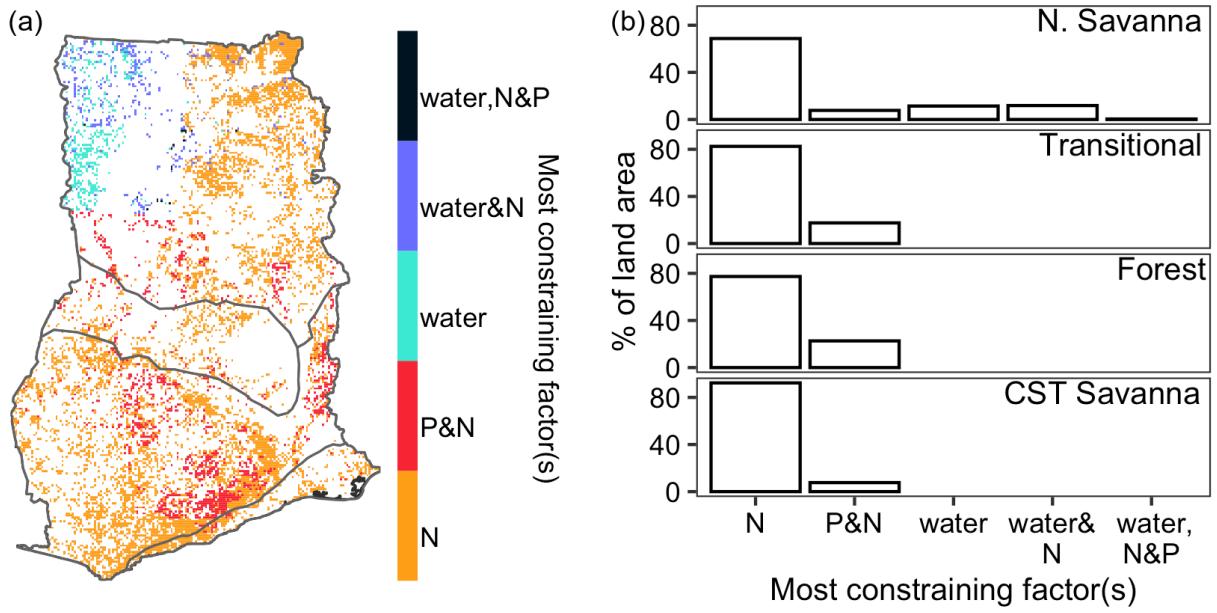

Figure 17 (a) Most yield-constraining factor(s) for rainfed unfertilized yam cultivation in Ghana, and (b) percentage of land area subject to the yield constraining factor(s) for the four agroecological zones. Abbreviations in (b): N. Savanna – northern savanna, and CST savanna – coastal savanna.

The tested improved management treatments reduced nutrient stress but not water deficiency stress. Water deficiency stress was slightly increased (by less than 7% across all simulation units) under the improved management treatments compared to the control treatment (Figure 18). On average, pigeonpea-yam rotation affected water stress most among the three improved management treatments. The rotation treatment resulted in about 2-3% increase in drought stress on average compared to the control treatment for the four agroecological zones, whereas about 1% or less increase was found under the pigeonpea residue incorporation and fertilizer addition treatments (Figure 18e-h).

Pigeonpea residue incorporation decreased N deficiency more than P deficiency. On average, the residue incorporation treatment reduced N stress level by about 20% for coastal savanna, northern savanna and transitional zones, and about 15% for the forest zone. P deficiency stress

was reduced by an average of 4% for the forest zone and 2% for the transitional zone under the residue incorporation treatment (Figure 18a-d). By contrast, P stress under the residue incorporation treatment was higher than the control treatment in some areas of the coastal and northern savanna zones, and lower in other areas. On average, residue incorporation had a minor effect on P stress levels (change of less than 0.2%) for the two savanna zones (Figure 18a-d).

The pigeonpea-yam rotation treatment reduced N and P stress levels compared to the rainfed unfertilized treatment. For forest and transitional zones, on average, N stress level decreased by 5% and P stress decreased by about 9% in comparison to the control treatment. For the coastal and northern savanna zones, N and P stress were reduced by about 8% and 6%, respectively (Figure 18e-h).

Synthetic N and P fertilizer addition reduced both N and P stress. N stress level decreased by 6-9% in the coastal savanna, northern savanna, and forest zones, and by 4% in the transitional zone. P stress was reduced by 10-14% in the forest, transitional, and northern savanna zones and by 7% in the coastal savanna zone (Figure 18i-l).

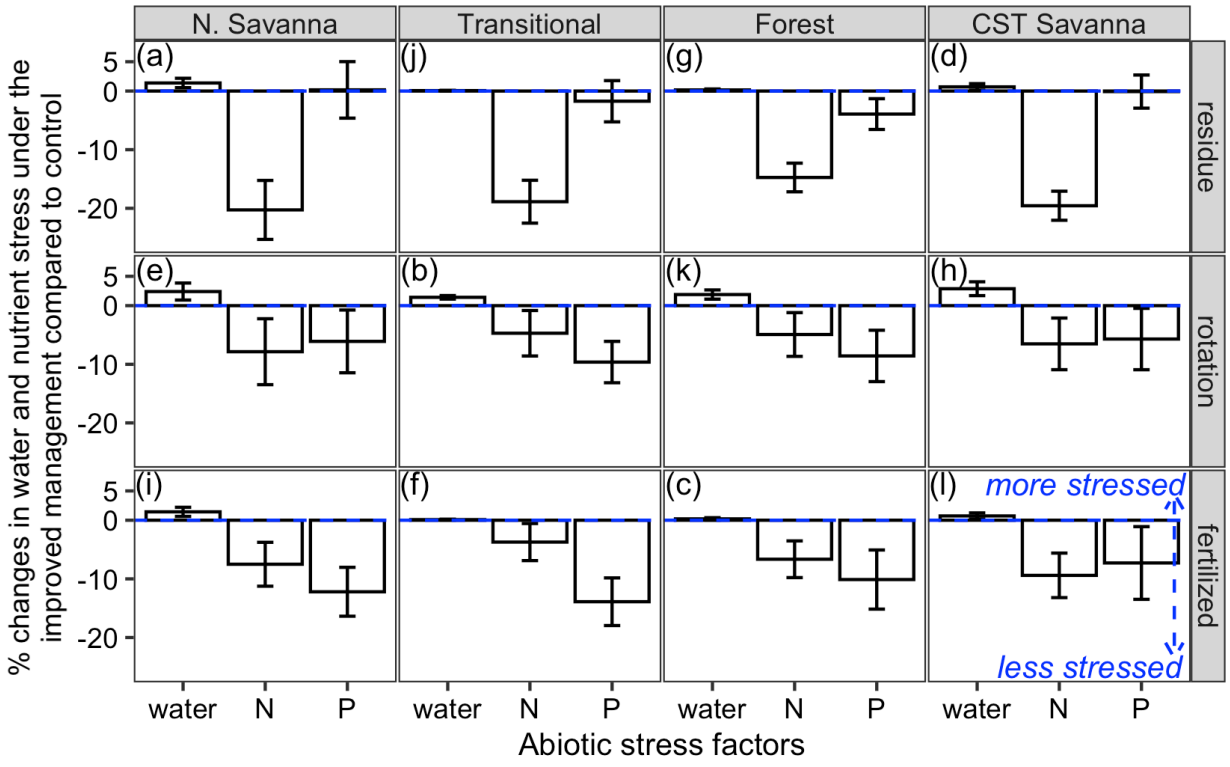

Figure 18 Average percentage changes in water, N, and P deficiency stress levels under each improved management treatment, including (a-d) pigeonpea residue incorporation, (e-h) pigeonpea-yam rotation, and (i-l) 23 kg/ha N-P_2O_5 fertilization, across the four agroecological zones (error bars represent standard deviation)

SOC Response to Management Treatment

SOC decreased at a rate ranging from 0.18 to 0.99 Mg/ha/year across Ghana over the simulated 10 years under unfertilized rainfed yam cultivation (Figure 19b). The forest zone lost the most SOC, at an average rate of 0.63 Mg/ha/year and standard deviation of 0.06 Mg/ha/year, followed by the transitional and coastal savanna zones (average SOC loss rate of 0.50 Mg/ha/year, respective standard deviation of 0.098 Mg/ha/year and 0.067 Mg/ha/year) and the northern savanna zone (average loss rate of 0.37 Mg/ha/year and standard deviation of 0.095 Mg/ha/year) (Figure 19a).

When pigeonpea residue was incorporated, SOC increased over 10 years by less than 0.25

Mg/ha in about 0.3% of cropland area—located in the northern savanna—but it continued to decline for the rest of the cropland in Ghana. SOC was lost at a slower rate than the control (rainfed unfertilized) treatment (Figure 19c). SOC in the forest zone decreased at an average rate of 0.43 Mg/ha/year, an average of about 0.30 Mg/ha/year for the transitional and coastal savanna zones, and an average of 0.16 Mg/ha/year for the northern savanna zone. The standard deviation of SOC loss for the four agroecological zones were 0.063 Mg/ha/year, 0.097 Mg/ha/year, 0.063 Mg/ha/year, and 0.094 Mg/ha/year, respectively (Figure 19a).

Compared to the control treatment, SOC decreased less under the pigeonpea-yam rotation and fertilizer addition treatments. Under the two improved management treatments, SOC loss rate ranged from 0.09 to 0.93 Mg/ha/year (Figure 19d-e). Under the two management treatments, SOC loss was greatest in the forest zone (an average of 0.55 Mg/ha/year), followed by the transitional and coastal savanna zones (average about 0.4 Mg/ha/year) and the northern savanna zone (average of about 0.3 Mg/ha/year). The standard deviation of SOC loss for the forest zone under the two management treatments was about 0.07 Mg/ha/year, 0.1 Mg/ha/year for the transitional zone, 0.07 Mg/ha/year for the coastal savanna zone, and 0.09 Mg/ha/year for the northern savanna zone (Figure 19a).

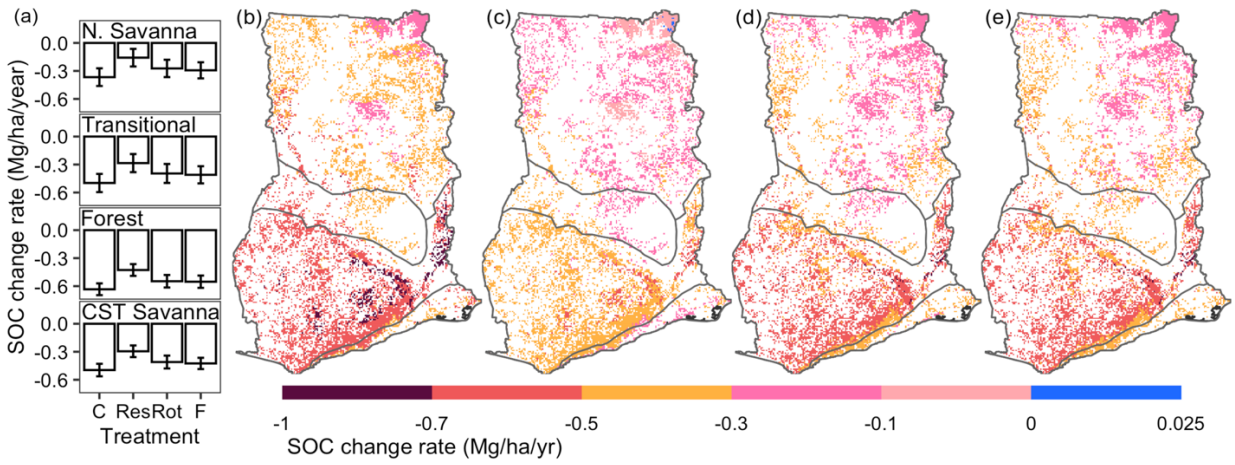

Figure 19 Response of soil organic carbon (SOC) change rate to management treatments across Ghana. (a) average SOC change rate by agroecological zone under the four treatments, and (b-e) SOC change rate under the four respective treatments: control, pigeonpea residue incorporation, pigeonpea-yam rotation, and yam with fertilizer addition. Abbreviations in (a): N. Savanna – northern savanna, CST savanna – coastal savanna, C – control, Res – residue, Rot – rotation, and F – fertilized. Error bars in (a) represent standard deviation.

Discussion

Accuracy in Simulating Yam Yield Response to Management Treatments

Unlike grain crops (e.g. maize and wheat), yam plants have not been extensively evaluated for crop modeling (Raymundo et al., 2014). While other tuber-based crop models (e.g. CropSystVB-Yam) explicitly simulate tuber induction and bulking processes, our SALUS-Yam model is similar to the EPIC-Yam model in that it uses generic crop parameters to simulate yam biomass productivity (with considerations for soil water and nutrient dynamics), and uses harvest index to partition simulated biomass into aboveground (vine and leaf) and tuber categories (Marcos et al., 2011; Srivastava and Gaiser, 2010). In addition to accurately simulating biomass yield—both for tubers and aboveground—for six treatments (combining two yam-based systems, yam monoculture and yam with pigeonpea residue incorporated, with three fertilizer rates) at two locations for two years, our simulated tuber yield under the tested management matched with the reported yield in the literature. Our simulated rainfed unfertilized tuber yield in 2016 (6.9 Mg/ha)

matched with the reported national rainfed tuber average yield of 6.1 Mg/ha, given 65% moisture content (Frossard et al., 2017; MoFA, 2017). The simulated tuber yield under rainfed and unfertilized conditions in the transitional zone (3-7 Mg/ha) also matched with yield observed in farmer's fields (about 2 Mg/ha) and experimental station trials (5 Mg/ha) in Ghana (Danquah et al., 2018; Ennin et al., 2014). Yam tuber yield response to fertilizer addition agreed with the reported values in the literature. With 23 kgN/ha and 10 kgP/ha of fertilizer added, tuber yield increased by 19-89%, and an average of 44%. Similar yam yield response to N and P fertilizer input was observed in field experiments in Benin and Nigeria. In Benin, yam yield increased by 30% with 30 kg/ha of N and P fertilizer compared to unfertilized yield (Srivastava et al., 2010). In Nigeria, yield increased by 53-71% and 71-84% with 15kg/ha and 30 kg/ha N-P2O5 fertilizer addition, respectively (Law-Ogbomo and Egharevba, 2009; Law-Ogbomo and Remison, 2008).

Accuracy in Simulating SOC Response to Management Treatments

SOC dynamics are impacted by land use, initial carbon stock, weather conditions, soil texture, and nutrient return to the soil (Bruun et al., 2015; Liu and Basso, 2020a; Owusu et al., 2020). Our results revealed that SOC declined after 10 years of yam cultivation for more than 99% of cropland in Ghana, regardless of management regime. Our simulated SOC change was comparable to the reported SOC decline rate in sub-equatorial forest and Sudano-Sahelian savanna regions in Ivory Coast. In our study, the SOC decline rate was about 1 Mg/ha/year in some areas of the forest zone, where total annual precipitation was about 1,250 mm. Similarly, the reported SOC decline rate was 1.9 Mg/ha/year for cereal cultivation in sub-equatorial forest in Ivory Coast with the annual precipitation of 2,100 mm (Bationo et al., 2007). Our simulated average SOC decline rate in the Sudan savanna zone was 0.15 Mg/ha/year, compared to the reported 0.16 Mg/ha/year under cereal cultivation Sudano-Sahelian savanna regions in Ivory

Coast (Bationo et al., 2007). Our average SOC loss rate of 1.28% per year under continuous unfertilized yam cultivation was comparable to the reported SOC loss rate of 1.5% per year with unfertilized sorghum cultivation in Burkina Faso (Bationo et al., 2007). We found that after 10 years of yam cultivation, SOC was higher under the three improved management treatments—residue incorporation, pigeon-pea-yam rotation, and fertilizer addition—than under control management treatment (on average 6.0%, 2.6%, and 2.2% higher, respectively). This finding is consistent with the slower SOC loss rate under improved management (rotation, fertilizer application, and both fertilizer and residue addition) of millet-groundnut compared to unfertilized treatment in Senegal. It conflicts, however, with studies of sorghum in Burkina Faso in which the reported SOC loss rate was higher with fertilizer application than without (Bationo et al., 2007). Our simulated SOC decline rate was also comparable to that observed in a 25-year millet-cowpea rotation experiment in Niger. With 3 Mg/ha residue incorporated, our simulated SOC declined at a rate of 0.013 Mg/ha/year in low-SOC-stock regions of the northern savanna. By comparison, with residue incorporated, SOC declined by about 0.02 Mg/ha/year under millet-cowpea rotation with residue addition (Nakamura et al., 2011). Our SOC decline rates under rainfed unfertilized (average loss rate of 0.51 Mg/ha/year) and fertilized (0.43 Mg/ha/year) treatments were higher than that in the 25-year millet-cowpea rotation field experiment in Niger (0.02 Mg/ha/year and 0.048 Mg/ha/year under unfertilized and fertilized treatment, respectively) (Nakamura et al., 2011). In addition, our results suggested that with high residue input, SOC could be sustained or increased slightly. This finding agrees with the observed increase in SOC with residue incorporation plus fertilizer addition for 25-year millet-cowpea cultivation in Niger (Nakamura et al., 2011).

Yam Yield-Limiting Factors

While uncertainties in rainfall onset and amount continue to be a challenge for yam production in Ghana, our results indicated that nutrient deficiency remained the most limiting factor for yam production for most of Ghana, northern Ghana included. Further, our results showed that increased yam tuber yield under improved management was due to N and P stress relief. Our results align with findings in the literature that managing nutrients is essential for Ghana's agriculture (Bationo et al., 2018; Van der Velde et al., 2014). Our results also support other research efforts to promote legumes, particularly pigeonpea, as a resource for sustainable yam production in Ghana. In addition to providing staking for yam cultivation, pigeonpea adds N to soil through biological N fixation and returns quality residue with low C:N ratio to soils (Acheampong et al., 2019; Ennin et al., 2014; Naab et al., 2015). Our simulation results, however, suggested that compared to the conventional unfertilized rainfed management, water deficiency stress could increase under the tested improved management treatments. This is due to larger yam plants, increased plant water use, and the consequent reduced soil water supply under improved management treatments.

Simulation Uncertainties

Despite high confidence in our simulations, soil and weather inputs were two major sources of uncertainty. The soil input data was derived from the AfSIS database, which provides gridded soil profile information at two spatial resolutions (1km and 250m) and three estimates (high, medium, and low values). In this study, we used the coarser dataset (1km resolution) with medium estimates for two reasons. First, we found the estimated organic carbon from the finer-resolution soil dataset (250m resolution with only medium estimates available) was unreasonably

high (the reported average organic carbon in the top 30cm of soil ranged from 0.2% to 5.5%, averaging 1.02%) compared to the medium estimate in the 1km-resolution dataset (0.4-2.2% with an average value of 1.07% in the top 30cm). We did not use the lower organic carbon estimate due to its much lower values, ranging from 0.03% to 0.6% with an average value of 0.24% across Ghana. Weather input data came from the gridded 0.5°-resolution POWER dataset. Gridded climate reanalysis products based on weather stations and remote-sensed precipitation events may overestimate precipitation through the inclusion of so-called drizzling precipitation events of less than 1mm (Sun et al., 2006; Valdés-Pineda et al., 2016). We, however, did not find systematic under- or over-estimation of rainfall when comparing POWER total annual precipitation to the total precipitation recorded in 20 weather stations throughout Ghana (APPENDIX C Figure 37a). Nor did we find substantial differences in simulated tuber yield between the original NASA POWER precipitation series versus a drizzle-precipitation-events-removed series (differences within 8%, and within 4% for more than 97% of cases, APPENDIX C Figure 37b).

Conclusions

Proper agricultural management is key to increasing crop productivity, conserving the environment, and improving the well-being of smallholder farmers. Yam is mostly produced in central and northern Ghana, where soil fertility is lower and rainfall is less frequent than in the south of the country. Shifting agriculture, in which crops are planted on clear-cut fields until soil nutrients are exhausted, is unsustainable. Yet, it is a common practice for growing yam. Yam management in smallholder fields in Ghana needs to not only increase short-term yield but also enhance long-term soil fertility. Many Ghanaian farmers are aware of improved management

strategies for enhancing yam yield, such as pigeonpea-yam rotation, mulch or residue application, and the use of synthetic fertilizer. Tuber yield and SOC benefits from the practices, however, have not been studied systematically. Our study investigated 10-year yam yield and SOC changes under both the commonly practiced rainfed unfertilized management and three improved practices, which included continuous yam cultivation with pigeonpea residue incorporation, pigeonpea-yam rotation and fertilized continuous yam cultivation, through the lens of crop modeling that explicitly considered the varying levels of soil fertility and rainfall abundance across Ghana. This study sheds light on yam response to pigeonpea residue application, rotation with pigeonpea, and synthetic fertilizer application across Ghana, considering varying levels of soil fertility and different rainfall patterns. We provided a benchmark study to evaluate yam yield and SOC response to various management practices across Ghana and make yam management decisions. Future research needs to explore the effects of initial soil nutrient stock, land cover, and cropping system on crop yield and SOC.

Acknowledgements

This work was made possible through support provided by the US Agency for International Development (USAID) under the terms of Cooperative Agreement No. 7200AA18LE00003. We would like to thank the staff members at Council for Scientific and Industrial Research-Crops Research Institution in Ghana for maintaining the field experiments and collecting the data that was used to calibrate and validate the model. We would like to thank the Ghana Meteorological Agency for sharing weather observations recorded by weather stations across Ghana.

CHAPTER 5: ASSESSMENT OF FIELD VARIABILITY IN SMALLHOLDER FIELDS USING SATELLITE IMAGES

Abstract

The main goal of this paper was to address an important, yet infrequently discussed aspect of smallholder farming, within-field variability. The objectives were to assess within-field variability in smallholder fields in Sub-Saharan African countries, and to evaluate the impact of spatial resolution of satellite images in monitoring vegetation status. I used field observations of in-season plant density and yield at harvest, and different-resolution remote sensing images to evaluate within-field variability in maize-based fields in Tanzania, yam-pigeonpea fields in Ghana, bean growing areas in Honduras and large commercial farms in the US. Within-field variability was quantified using coefficient of variation (CV). Across the study sites in three districts in Tanzania in the long and short rain seasons in 2017-2018, maize-based fields had much within-field variation regarding plant condition, plant density, and grain yield. The CV of within-field plant density was 7.2-80.2%, and the median CV of within-field plant density was about 20-30% for the three districts, Morogoro, Kagera, and Tanga. At the end of growing seasons, CV of grain yield ranged from 9.7% to 93.8%, with median values of 31.3%, 36.3% and 29.7% for the three districts, respectively. The grain yield variability was significantly correlated with any of the four in-season vegetation indices — Normalized Difference Vegetation Index (NDVI), green NDVI, Green Chlorophyll Vegetation Index (GCVI), and Enhanced Vegetation Index (EVI) — derived from the PlanetLab's PlanetScope 3m-resolution images taken during the critical maize growth period (about 1-2 months after planting; $p < 0.1$, $p < 0.05$ or $p < 0.01$ depending on the index) for fields in Morogoro. Among the vegetation indices, within-field

GCVI variability had the strongest correlation with within-field yield variability (Spearman's correlation coefficient: -0.7). Regarding spatial resolution's impact on monitoring in-season crop growth, the in-season within-field satellite-image-derived NDVI variability decreased with coarser spatial resolution of satellite images, from 3m PlanetScope, to 10m Sentinel-2, to 30m Landsat. This was true for smallholder yam and pigeonpea fields in Ghana (N = 9) and large fields in the US (N = 6). This trend also held for 10km x10km bean growing areas in Honduras (N = 7). However, such differences were insignificant ($p > 0.05$). This work showed considerable within-field variability in smallholder fields and the possibility of detecting such variability using high-resolution satellite images. This work also implied that high-resolution remotely sensed images could contribute to improving land resource management for smallholder fields. More investigations on the compounding effect of spatial and spectral resolution of different sensors are needed to assess the capability of remote-sensing images for monitoring field variability.

Introduction

Remote sensing technologies have been transforming the agricultural sector by offering large spatial and temporal coverage of land area. Continuous monitoring of vegetative areas progressing from pre-planting, through mid-season, maturity, and harvest, to the next growth cycle provides insights on greenness development and possible abiotic (e.g. N and water deficiency) and biotic (e.g. fungi and pest infection) stress that crops experience (Liaghat and Balasundram, 2010; Mulla, 2013). The use of remotely sensed images — acquired from satellites, unmanned aerial vehicles (UAV, such as drones), and handheld devices —in precision N management is one example of integrating remote sensing technologies for better management

of agricultural resources (Basso et al., 2016a; Duan et al., 2017; Teal et al., 2006). Such advances have taken places mostly in the US and European countries, and recently in India and China. Smallholder farmers in African countries and other low- or low-to-middle- income countries are still waiting for such advances to transform how they manage their land.

Applying remotely sensed images to managing smallholder fields poses multiple challenges. In addition to cloud and haze contamination and sensor calibration, which are common sources of error for satellite image analysis, the unique features of smallholder fields — uneven field boundaries, fragmented landscape, and diverse crops in the field — present an extra layer of complexity for inferring agronomic information from remote-sensing images (Jain, 2020). Recent work on applying satellite or UAV images to smallholder agriculture has largely focused on crop yield prediction (Burke and Lobell, 2017; Iizumi et al., 2018; Yonah et al., 2018), along with a few studies on detecting land cover (Sibanda and Murwira, 2012; Sweeney et al., 2015), field boundary (Persello et al., 2019), and phenology (Duncan et al., 2015).

One critical, but often ignored, aspect of managing agricultural fields is coping with within-field crop variations. Large fields are highly variable due to soil properties and landscape positions, as has been shown in numerous studies (e.g. Godwin and Miller, 2003; Maestrini and Basso, 2018a; Robertson et al., 2008). Field variability, nonetheless, also exists for smallholder fields. It requires on-the-ground agronomic information to translate light reflectance information captured by remotely sensed images to tangible management action. Much work awaits to reveal within-field variability and the ability of the current very high resolution satellite images to capture variations within a small field.

Here, I asked two different research questions: (1) To what extent can high-resolution images predict within-field yield variability for smallholder maize fields in Tanzania? (2) How much

field heterogeneity information is lost due to resolution of remote sensing images? I demonstrated within-field variability in smallholder fields, and investigated the capability of satellite images to detect such field variability. The objectives were to assess within-field variability in smallholder fields in Sub-Saharan African countries, and to evaluate the impact of spatial resolution of satellite images for monitoring vegetation status.

Materials and Methods

Description and Detection of within-Field Variability in Maize Fields in Tanzania

I reported within-field variability regarding maize plant density and final yield. I supplied photos taken during the field campaign to further illustrate within-field variability in smallholder fields in Tanzania. For each field, one or two 6x6m sampling plot(s) were established at a random location within a field. Each sampling plot was divided equally into four 3x3m subplots. The variability, expressed in coefficient of variation (CV), was calculated across the subplots for each field.

To assess the potential of detecting yield variability prior to harvest by satellite images, I used Spearman correlation coefficient between within-field yield variability and satellite-image-derived vegetation index for each of available satellite images. I used PlanetLab's very high resolution PlanetScope Ortho Tile multispectral images (3m spatial resolution and almost daily temporal resolution) (downloaded from https://www.planet.com/products/planet-imagery/). The correlation analysis was conducted in a 10-day moving window when at least six field-scene combinations were available for each district due to different planting dates in the three districts (Morogoro was in a long rain season whereas the other two districts were in a short rain season). Scenes within 2 months after planting in 2017 with less than 30% cloud coverage from

PlanetScope were downloaded and processed for four vegetation indices, Normalized Difference Vegetation Index (NDVI), green Normalized Difference Vegetation Index (gNDVI) Green Chlorophyll Vegetation Index (GCVI), Enhanced Vegetation Index (EVI) (Equation 5-8). Cloud pixels (NDVI < 0.05) was removed before the analysis.

$NDVI = (NIR-Red)/(NIR + Red)$ ……..…Equation (5),

$gNDVI = (NIR-Green)/(NIR + Green)$ ……Equation (6),

$GCVI = (NIR/Green) - 1$…….……..……Equation (7),

$EVI = 2.5*(NIR – Red)/(NIR + 6*Red – 7 * Blue + 1)$ ……...Equation (8),

Where NIR, Red, Green and Blue means reflectance from near infrared (NIR), red, green and blue band, respectively.

Impact of Remote-Sensing Image Spatial Resolution on Monitoring Field Variability

To answer the second question, I surveyed possible study areas in three countries: (1) Honduras, where crop cultivation is constrained by precipitation gradient and beans are of interest for food security issues, (2) Ghana, where yam is a major food crop, and (3) the US, where soybean, maize and wheat are important commodity crops. Crop cultivation takes place mostly in smallholder fields in Honduras and Ghana, whereas it occurs in large commercial farms in the US.

For Honduras, I first identified growing areas of beans based on a global crop production dataset, Spatial Production Allocation Model (SPAM), prepared by the International Food Policy Research Institute (IFPRI, 2019; You et al., 2009). I first selected 8 areas (10x10km each) within Honduras where the beans harvested areas were above 700 ha in 2010 (APPENDIX D Figure

39). I then classified the land cover for the 8 bean growing areas using a random forest classifier supervised classification method on the Google Earth Engine to refine agricultural pixels within each bean growing area. I chose beans as a study crop because they were the most extensively grown legume crop in Honduras (You et al., 2009).

For Ghana, I selected 12 yam and pigeon-pea fields occupying an area of 100m X 90m in Fumesua, Ghana (6°41' N, 1°28'W). The 12 fields consisted of four distinct cropping systems, sole pigeonpea, sole yam, pigeonpea-yam intercropping and, yam with 2-3 rows of pigeonpea at the border. Each cropping systems contained three replicates. The size of each field ranged from 244 m^2 to 687 m^2 with an average of 476 m^2 (APPENDIX D Figure 40). In each field, three fertilizer treatments (45-45-60, 23-23-30, and 0-0-0 N-P2O5-K2O kg/ha) were arranged in a split plot design.

For the US, five farms in 2017-2019 in Ionia county of Michigan were selected (42°52' N, 84°55' W). A total of 6 field-year combinations were included for this study. Maize was planted in two fields in 2017 and two fields in 2018. Soybean was planted in one field in 2018 and wheat was planted in one field in 2019. The size of the five unique fields ranged from 15.3 ha to 34.4 ha, with an average of 26.8 ha and a median of 27.6 ha.

I selected three accessible satellite imageries that have distinct spatial and temporal resolution, the almost daily very fine resolution PlanetScope (3 m), fine resolution at a 10-day interval Sentinel-2 (10 m) and moderate resolution every 16 days Landsat-8 or Landsat-7 (Landsat) (30 m). In addition, UAV images, ranging from less than 10cm to 1m resolution, were available for maize and wheat fields in the US, and were used to further illustrate the impact of spatial resolution on in-season vegetation variability. The spatial resolution of images from the four remote sensing platforms were illustrated in Figure 20. For the satellite images, I

downloaded scenes from Landsat-8 (supplied with Landsat-7 when no available scenes were available), Sentinel-2 and PlanetScope with less than 30% cloud coverage during crop's critical growing window and removed cloud pixels of the scenes for each study area in the three countries. Crops' critical growing window is before the start of reproductive stage in which harvestable yield is highly sensitive to environmental stressors and in which satellite-derived vegetation index may explain the final yield or yield variations (from results for research question 1 in this paper, Liu and Basso, 2020a; Maestrini and Basso, 2018b). I selected scenes in which peak NDVI was observed in November in 2018 for beans in Honduras, early June in 2018 for yam in Ghana, and early-to-mid July for maize, mid August for soybean and early June for wheat in the US, considering precipitation patterns, the critical time window and available cloud-free scenes overlapping similar dates (within about one week difference) (Table 5, APPENDIX D Figure 41). I chose within-field NDVI variation as a proxy for variability. For Ghana, only PlanetScope and Sentinel-2 images were used because a Landsat 30m scene covers more than one study field and is inadequate to reveal within-field variability. For the UAV images of maize fields, the spatial resolution was 0.08-0.43m, average of 0.2m. One UAV scene was available per field for two maize fields in July and one wheat field in June (Table 5). NDVI of one maize field in the US derived from the three satellite images and UAV were shown in Figure 21. I compared the CV of the skewed-distributed NDVI across the three satellite imageries using the Kruskal-Wallis test for sites in each country at $P = 0.05$.

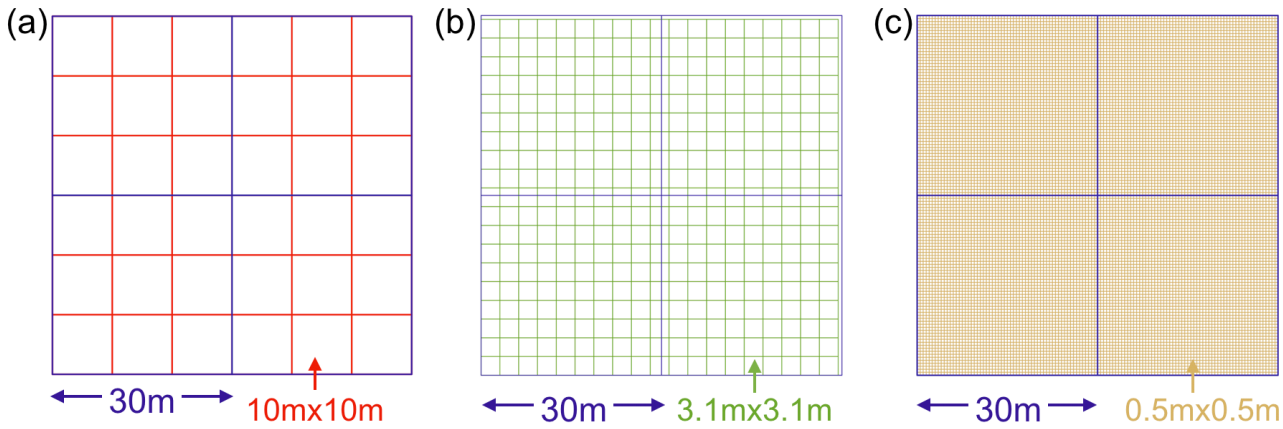

Figure 20 Illustration of spatial resolution of three satellite images and UAV images. (a) 4 pixels of 30m resolution Landsat image grids versus 36 pixels of 10m resolution Sentinel-2 image grids, (b) 4 pixels of 30m resolution Landsat image grids versus 361 pixels of 3m resolution PlanetScope image grids, and (c) 4 pixels of 30m resolution Landsat image grids versus 90000 pixels of 0.5m resolution UAV image grids (Landsat grids were outlined in blue, Sentinel-2 grids outlined in red, PlanetScope grids outlined in green, UAV grids outlined in light brown)

Table 5 Summary of satellite scenes for study fields in Honduras, Ghana and the US

PlanetScope	Sentinel-2	Landsat-8*
	beans in Honduras	
----------	Peak NDVI dates in November in 2018	----------
	yam and pigeonpea in Ghana	
Jun. 4, 2018	Jun. 6, 2018	N/A
	maize in US	
----------	Early-to-mid Jul., 2017-2018	----------
	soybean in US	
Aug. 11, 2018	Aug. 11, 2018	Aug. 19, 2018*
	wheat in US	
Jun. 3, 2019	Jun. 7, 2019	Jun. 3, 2019

*Landsat-7 image was supplied when cloud-free scenes from Landsat-8 was not available

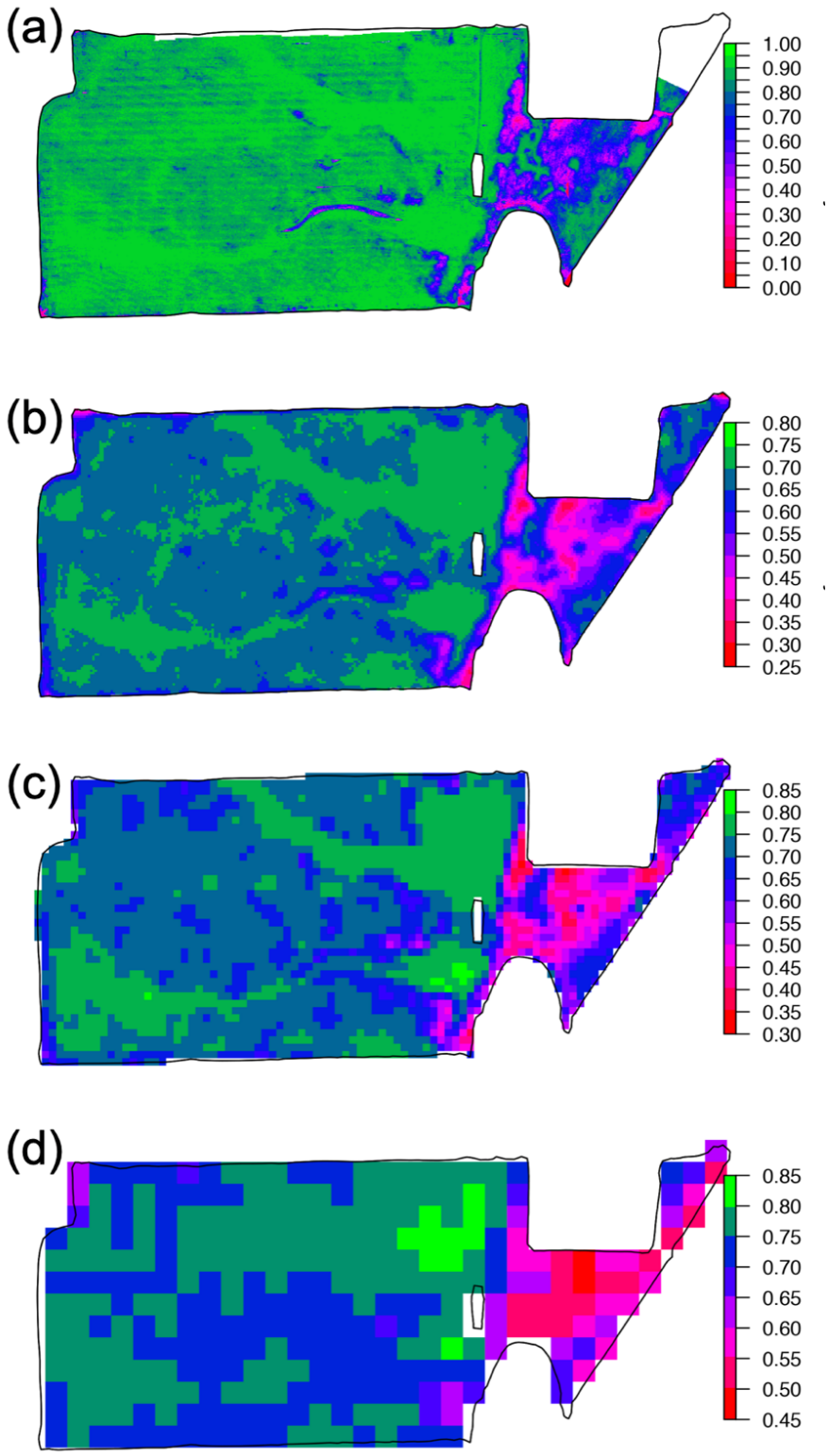

Figure 21 NDVI derived from a (a) UAV image (0.12m), (b) PlanetScope image (3m), (c) Sentinel-2 image (10m), and (d) Landat-7 image (30m) on July 10, 2018 for a maize field in Michigan, US

Results

Description and detection of within-field (yield) variability in maize fields in Tanzania

Though more than 90% of the study sites across Tanzania were less than 2ha in area, various levels of within-field variability exhibited. During growing season, in some fields, maize in the two sampling plots within a field looked alike (i.e. low variability), whereas the mid-season maize was in considerably different condition in other fields (i.e. high variability) (Figure 22a-b). Maize plant density in different locations of a field appeared to be different. The CV of plant density was 7.2-61.0% (median of 19.1%), 10.4-48.2% (median of 28.3%) and 10.9-80.2% (median of 25.3%) across the study sites in Morogoro, Kagera and Tanga (Figure 23a). At harvest, yield varied within a field. The CV ranged from 9.7% to 93.8% across the three districts, with respective median value of 31.3%, 36.3% and 29.7%. (Figure 23b). The within-field yield variability tended to increase with lower average yield (APPENDIX D Figure 42).

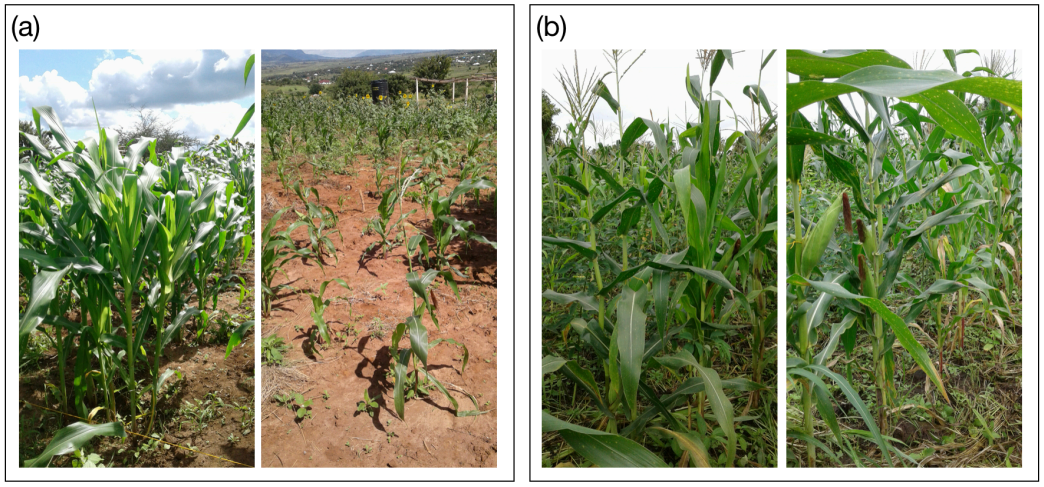

Figure 22 Photos from two sampling plots within a maize field with (a) high within-field variability and (b) low within-field variability

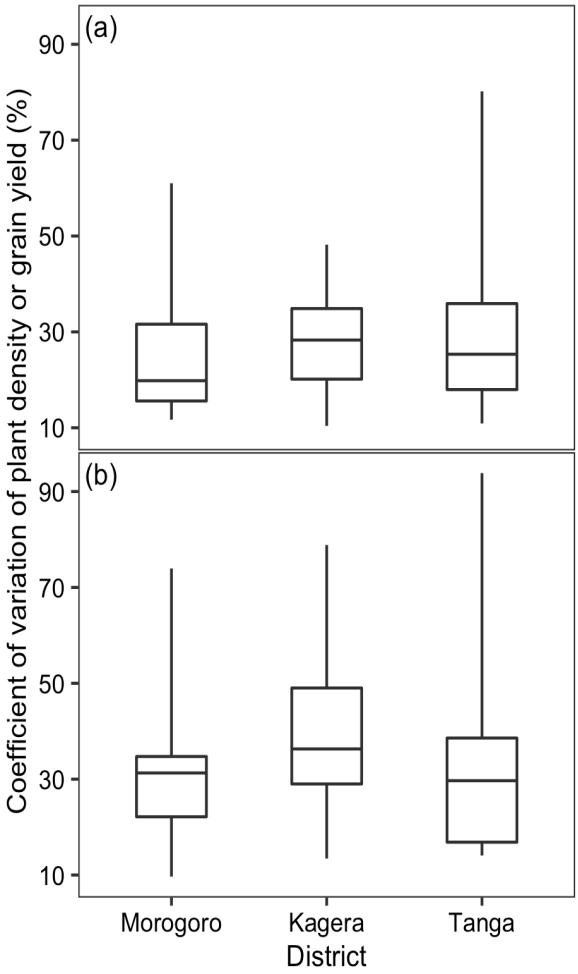

Figure 23 CV of within-field (a) in-season maize plant density and (b) grain yield at harvest across study sites in three districts in Tanzania

In general, correlation between yield variability (CV of within-field yield) and GCVI variability (CV of within-field GCVI) was the strongest, among the four evaluated vegetation indices from PlanetLab 3m resolution images. For Morogoro, the variability captured by remote sensing images in May 20-29, 2017 could explain final yield variability (Spearman correlation coefficient -0.70 to -0.49) with statistical significance ($p < 0.01$ for GCVI, $p < 0.05$ for EVI, and $p < 0.1$ for NDVI and green NDVI) (Figure 24a). The results showed a strong correlation between vegetation index variability and yield variability 1-2 months after planting for sites in Kagera (Spearman correlation coefficient about -0.5 for the four vegetation indices) but the

correlation was statistically insignificant ($p > 0.05$) (Figure 24b). Limited cloud-free scenes between planting and harvesting were available for Tanga. The strongest correlation (Spearman correlation coefficient about -0.35 for the four vegetation indices) was found in early season (in October), rather than late season (in late November-early December). The correlation was insignificant ($p > 0.05$) (Figure 24c).

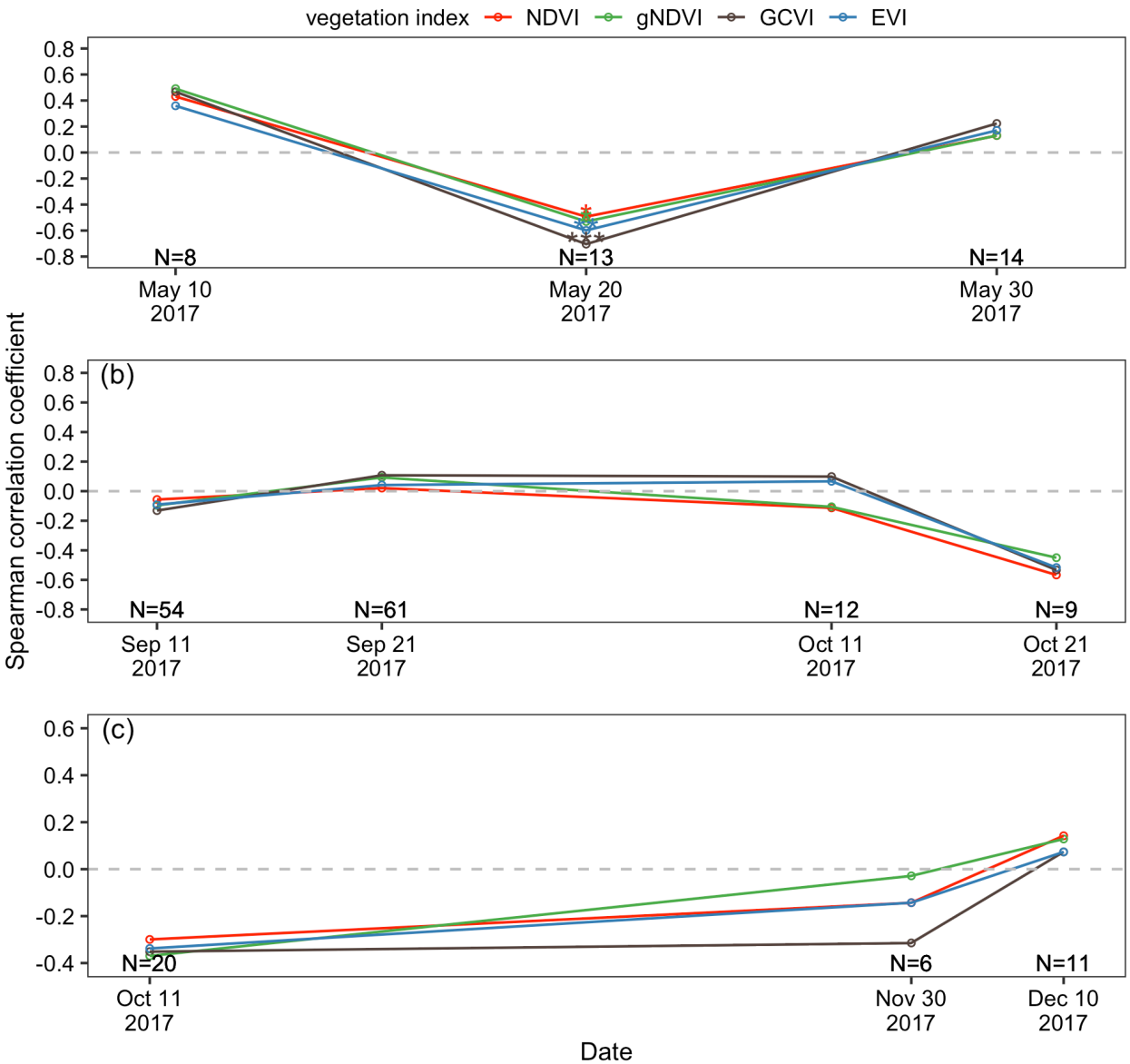

Figure 24 Evolution of Spearman correlation coefficient between within-field variability of vegetation index (four indices included) and within-field yield variability in maize growing season in 2017 for (a) Morogoro, (b) Kagera and (c) Tanga (the numbers in the bottom of each panel indicates the number of date-field combinations in the Spearman correlation analysis)

Impact of Satellite Image Spatial Resolution on Monitoring Field Variability

For Honduras, across seven bean-growing areas, where cloud-free scenes from three satellite images were found, the average within-field CV of NDVI was 18.4% according to 3m PlanetScope images, 18.3% using 10m Sentinel-2 images, and 17.4% using 30m Landsat-8 images. The respective standard deviations were 5.2%, 5.2%, and 6.5% (Figure 25).

For Ghana, the within-field NDVI variability ranged from 8.6% to 16.3% (average 11.3% and standard deviation of 4.3%) with the 3m-resolution PlanetScope and was 6.6-13.3% (average 10.3% and standard deviation of 2.9%) (Figure 26).

For the large commercial farms in the US, the CV of within-field 3m-resolution NDVI ranged from 3.9% to 10.7% (average 8.1% and standard deviation of 2.9%). Such variability was 1.7-14.0% (average 5.1% and standard deviation of 4.2%) with coarser 10m-resolution Sentinel-2 images, and was 2.3-12.9% (average 4.8% and standard deviation of 3.7%) based on 30m-resolution Landsat images. (Figure 27a). The CV of the 20cm-resolution UAV images for three fields was 8-13%, with an average of 10.1% and standard deviation of 2.8%. The results showed comparable NDVI variability detected by UAV, PlanetScope, and Sentinel-2 (CV about 10%) and much lower variability captured by Landsat (CV about 3.5%) (Figure 27b).

In general, the CV of NDVI increased with finer spatial resolution images for the selected study areas in Honduras, Ghana, and the US, but the difference was insignificant ($p > 0.05$) (Figure 25-27).

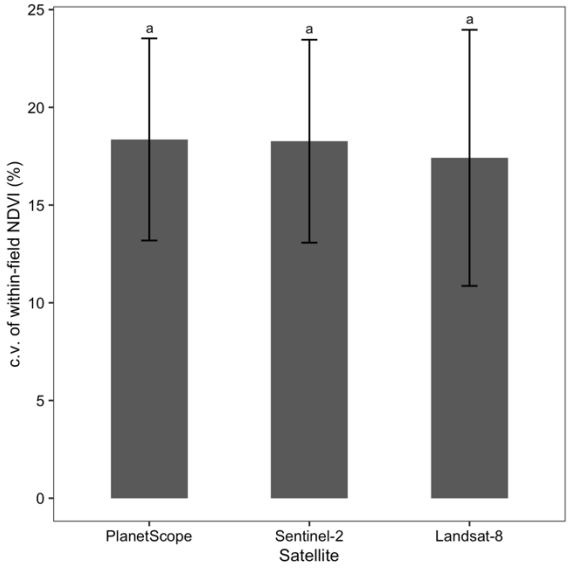

Figure 25 CV of within-bean-growing-area NDVI across the bean growing areas in Honduras (N = 7; the height of the bars indicates average value and the error bars indicate standard deviation; satellite images sharing the same letter are not significantly different ($p > 0.05$) using the Kruskal-Wallis test)

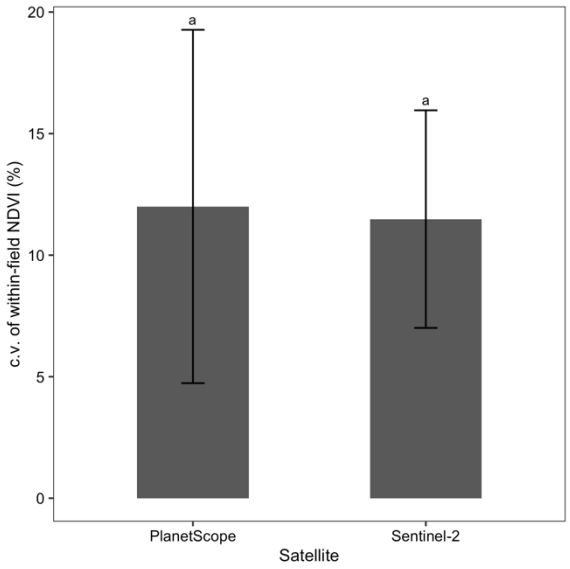

Figure 26 CV of within-field NDVI across yam-pigeonpea fields in Fumesua, Ghana (N = 9; the height of the bars indicates average value and the error bars indicate standard deviation; the same letter are not significantly different ($p > 0.05$) using the Kruskal-Wallis test)

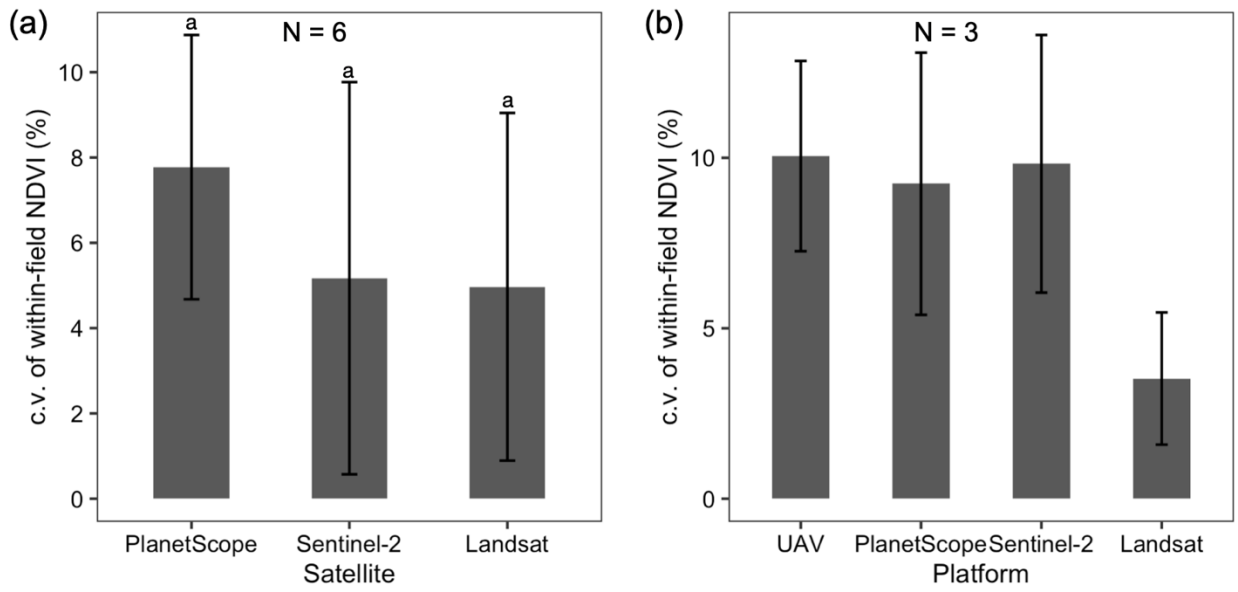

Figure 27 CV of within-field NDVI (a) from satellite images across six year-field combinations (N = 6), and (b) from both satellite and UAV images for three fields in Michigan, US. (N=3) (the height of the bar indicates average value and the error bar indicates standard deviation; satellite images sharing the same letter are not significantly different ($p > 0.05$) using the Kruskal-Wallis test)

Discussion

Smallholder fields contained various levels of within-field heterogeneity. In this study, I demonstrated such within-field heterogeneity using photos taken by tablets during field campaigns and vegetation indices extracted from very high resolution (3m) satellite images. I also quantified two aspects of such heterogeneity, in-season plant density and end-of-season yield, using field sampling data. In Tanzania, maize plant status may be drastically different depending on plant locations within a field. One location was filled with dense, thick maize plants during growing seasons while another location in the same field had sparse, short, and thin maize plants. This difference may be attributed to variations in soil properties within a field, as was indicated in the photos (Figure 22a). Tittonell et al. (2005) and Tittonell et al. (2007) showed a gradient of decreasing soil fertility and management intensity across fields within a farm (<1 ha) as they grew further from the homestead, and attributed within-farm variability to this

agronomic resource allocation gradient.

Assessing and managing within-field variability is critical to enhance crop yield. The traditional method of soil and plant sampling is labor and time intensive. The recent development of micro-satellites and ready-to-use satellite image products can aid in evaluating in-season within-field vegetation variability, as a proxy for end-of-season yield variability. For maize fields in Tanzania, I showed that grain yield variability was correlated to in-season (1-2 months after planting) PlanetScope-derived vegetation index (Spearman correlation coefficient above 0.5). One limitation of using vegetation index as an indicator for plant vigor and yield estimation for resource-limited small fields is that the detected greenness may be due to presence of weeds. A less varied vegetation index does not necessarily mean well-distributed crop growth within a field. This study was the first step in addressing within-field variability. I did not address management that could reduce such within-field variability. One study showed within-field yield variability across maize, sorghum, peanut, and cotton fields in Mali and its implications for fertilizer management (Schut et al., 2018). My work builds on such studies by voicing the need to evaluate variations in in-season plant conditions and final yield in order to effectively manage land resources and to improve crop yield for smallholder farmers.

One way to apply remote sensing information to manage field variability is to use the remotely sensed vegetation index to parameterize a model for management testing (Gilardelli et al., 2019; Jégo et al., 2012). But the capability of satellite images to detect vegetation variability varies depending on how temporal and spatial resolution relate to field size (Figure 25-27). This is particularly important for smallholder fields since their within-field variability has not been studied in relation to management planning. Though CV is useful to depict relative variability to a mean, its application is limited for assessing the impact of spatial resolution and sensor choice

on within-field vegetation variability detection. In this study, CV of within-field NDVI in general decreased with coarser sensor spatial resolution, but the impact of spatial resolution was statistically insignificant. This was partly due to varied NDVI ranges captured by different sensors with different spectral resolution (Figure 21). More research is needed to arrive at a conclusive understanding of the impact of compounded spatial and spectral resolution on vegetation status monitoring.

Conclusions

High variability, including variations in yield, plant status, and plant density, exists among different locations within smallholder fields in African countries. End-of-season yield variability can be detected by in-season PlanetScope (3m) images using derived vegetation indices, particularly GCVI (Spearman's correlation coefficient of -0.7). But use of satellite images for detecting within-field variability needs to be cautioned because temporal resolution of satellite images and the spatial resolution in relation to field size may limit its capacity.

Acknowledgements

This work was made possible through support provided by the US Agency for International Development (USAID) under the terms of Cooperative Agreement No. 7200AA18LE00003. I would like to thank Eric Owusu Danquah at Council for Scientific and Industrial Research-Crops Research Institution in Ghana for providing coordinates of pigeonpea and yam fields. The author was grateful for Rich Price for collecting UAV images, and Ruben Ulbrich and Sasmita Sahoo for pre-processing PlanetLab's imageries for the US fields. The author would like to thank PlanetLab for providing its micro-satellite images through its Education and Research Program.

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

This dissertation has presented the use of advanced agricultural technologies — including cloud-based data collected from tablets with field survey questionnaires coded in applications in portable devices (Chapter 2), process-based crop models (Chapter 3-4) and fine-spatial-resolution satellite images (Chapter 5) — in developing a timely and accurate maize yield forecasting system, assessing climate change impact on cropping systems, evaluating effects of agronomic management on yield and soil carbon, and assessing within-field variability for smallholder fields in Tanzania, Ghana, and Honduras.

Chapter 1 set the background and timely importance of this dissertation. It also included a brief overview of the process-based crop model, System Approach to Land Use Sustainability (SALUS), which was part of the methodologies for Chapters 2-4.

Chapter 2 presented a novel maize yield forecasting method that was able to predict maize yield in smallholder fields 14-77 days prior to harvest. The new method took advantage of available technologies — including computer-assisted personal interviewing software, tablets, and cloud information storage — to collect, store, and access timely in-season field information. The information from the field questionnaire survey was integrated with the SALUS crop model to provide accurate and timely yield forecasts.

Chapter 3 was centered around climate change uncertainties and their impact on maize fields in Tanzania. I concluded that under climate change, maize yield was most likely to decrease, particularly in Kagera (by average of 22-25% under RCP 4.5 and 8.5) and Tanga (by average of about 36% under the two RCPs), along with a more modest decrease in Morogoro (by an average of about 4%). Soil organic carbon was projected to decrease (by average of 22-25% across the three districts under RCP 4.5 and RCP 8.5), and soil inorganic N was most likely to increase (by

average of 1-7 kgN/ha across the three districts and two RCPs), in spite of large uncertainties in the projected climate. This study also highlighted the importance of considering changes in climate extremes in climate impact assessment, particularly for impacts on yield and soil inorganic N.

In Chapter 4, I investigated agronomic management treatments that could enhance both yam tuber yield and soil organic carbon in Ghana. The calibrated and validated SALUS-Yam model was executed to test the effectiveness of conventional and three improved management treatments (pigeonpea-yam rotation, yam with 3 Mg/ha pigeonpea residue incorporated, and yam with 23-23 kg/ha N-P2O5 fertilizer added) for increasing yam tuber yield and conserving soil carbon across all cropland in Ghana. This work provided a benchmark study to evaluate yam yield and soil carbon response to various management practices across Ghana and make yam management decisions. I concluded that among the tested agronomic practices, incorporating pigeonpea residue to yam fields would improve yield and reduce SOC loss to the largest extent.

Chapter 5 assessed the potential of using different satellite images, including 3m PlanetScope, 10m Sentinel-2, and 30m Landsat, in detecting within-field variability. Along with satellite images, field sampling data on in-season plant density and final grain yield across study sites in Tanzania was used to demonstrate sub-field-scale variability. I concluded that large variability existed among different locations within smallholder fields in African countries. In-season variability of Green Chlorophyll Vegetation Index, derived from PlanetScope images, was able to explain yield variability (r: -0.7). The capacity of satellite imaging for monitoring within-field variability, however, is limited by its spatial resolution in relation to field size.

With increasingly available and affordable technologies for agricultural digitalization, smallholder farming is facing a tipping point for transformation. The much-needed yield

improvement cannot be accomplished without knowing the factors that constrain its potential. The compounding effect of feedback between crop growth and soil nutrient cycling on soil fertility will need to be addressed for sustainable agriculture to succeed, particularly under climate change. Crop models can shed light on yield through crop-soil-climate-management interactions, and are effective in recommending site-specific management for better yield, as has been shown in previous and current research. Quality agro-climatic data remains a limitation for assessing and improving food crop yield for smallholder fields. While satellite images provide extensive spatial coverage for monitoring vegetation status, a lack of ground-truthing information, among other reasons such as cloud contamination, restricts their use for inference. The current fine-resolution micro-satellite images and computational algorithms may advance automatic delineation of smallholder fields and identification of field plants. Imaging small fields' within-field variability through unmanned aerial vehicles or very-fine-spatial-resolution satellite images can provide valuable soil and plant information for efficient management and thus yield improvement.

APPENDICES

APPENDIX A: Chapter 2 Supplemental Tables and Figures

We evaluated the capability of SALUS in simulating interannual yield variability at regional level by comparing the simulated yield to the reported regional yield from the Government of Tanzania in 2003/2004-2012/2013. We first ran the uncalibrated SALUS using possible management practices, a low-yielding maize cultivar for Morogoro, Kagera and Tanga with the respective typical soils in the regions and daily weather data in 2003-2013 extracted from National Aeronautics and Space Administration Prediction Of Worldwide Energy Resource (NASA POWER) (https://power.larc.nasa.gov). The possible management practices combined different planting dates (between late February and mid-April for long-rain seasons, and between late September and mid-November for short-rain seasons), nitrogen application rates (0, 30 and 60 kgN/ha), and planting densities (1-6 plants/m^2). A total of 648 simulations were performed for each district. We then fixed the planting date (early March for long-rain seasons, and early October for short-rain seasons) and selected N application rate and planting densities in each year that could reproduce the regional-level yield. We performed simulation runs with the management practices that would reproduce the reported yield at regional level, combined with typical soils for each region. We applied a 20% reduction factor to count for abiotic stress from weeds, pests and diseases (Tollenaar et al. 1994).

Without calibration, the range of simulated grain yield captured the reported yield at district level (Figure 28a). The average simulated grain yield using the uncalibrated SALUS model was able to capture interannual grain yield for Morogoro and Tanga reasonably well for 5 years of the simulated 10 years. On average, the uncalibrated model, however, overestimated grain yield in Kagera. For the three districts, the root mean square of deviation (RMSD) between the simulated average and reported regional yield was 0.75 t/ha with mean absolute percentage error (MAPE)

of 64.7% (Figure 28a). With selections of planting date and density for the model runs, the simulated grain yield was much closer to the reported regional yield across the three districts. Across the three districts, the RMSD between the simulated average yield and the reported yield was 0.23 t/ha, and the MAPE was 18.1% (Figure 28b).

The SALUS model also captured the yield response to stress levels (nitrogen deficiency stress and drought stress). The simulated stress factor in the SALUS model represents the ratio between supply and demand. The stress factor ranges from 0 (severe stress) and 1 (no stress) (Liu and Basso 2017). Using a constant planting density (5 plants/m^2) under rainfed and unfertilized condition, the simulated grain yield in general increased as the stress levels decreased (Figure 29).

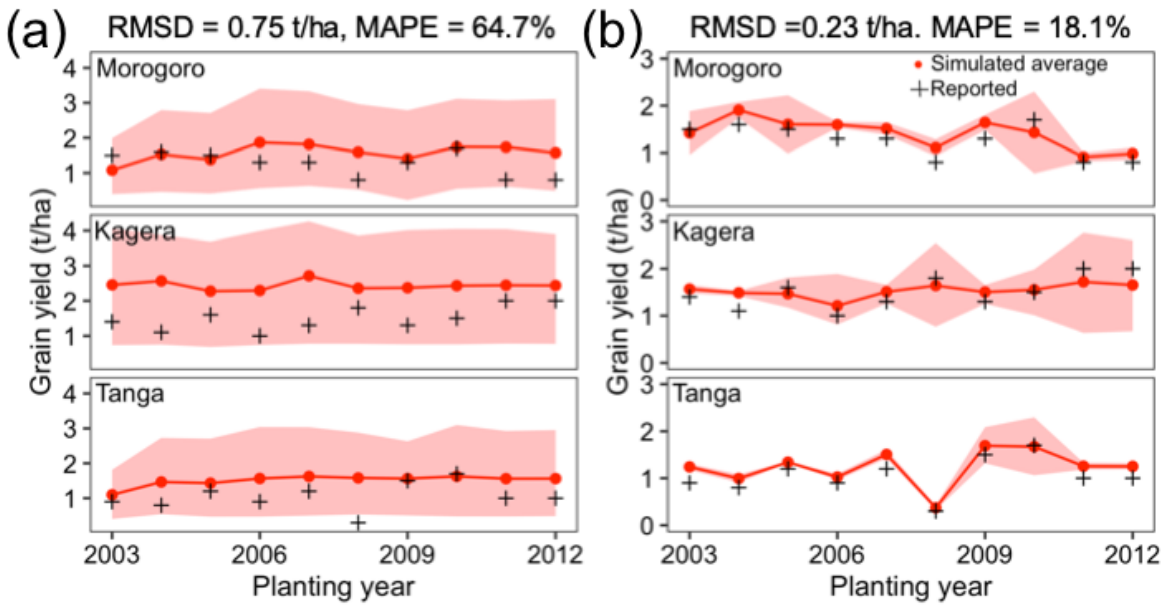

Figure 28 Comparisons between simulated yield and reported regional yield. (a) Simulated yield with all possible management practices versus reported regional yield, and (b) Simulated yield with management practices that would reproduce the interannual yield variability versus regional yield (In both panels, the black cross represents reported regional yield, the red dot represents average simulated yield, the upper and the lower boundary of the pink shades represent 90th percentile and 10th percentile values of the simulated yield, respectively; Abbreviations: RMSD - root mean square of deviation, MAPE- mean absolute percentage error).

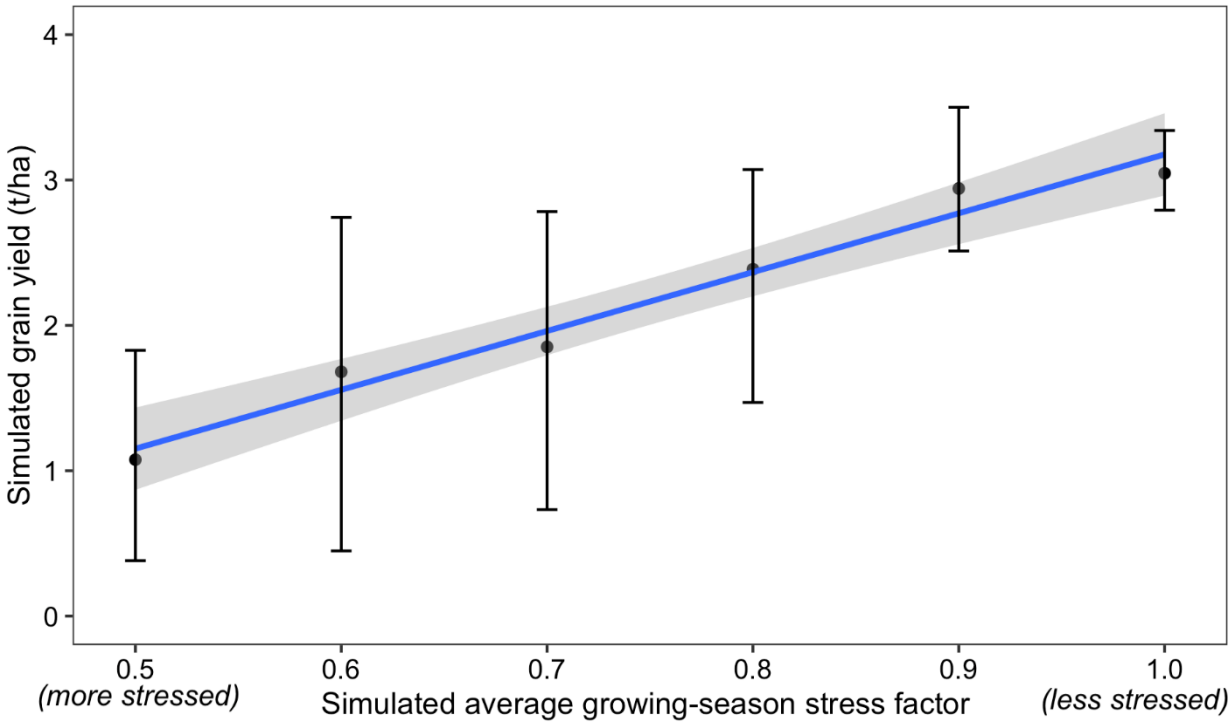

Figure 29 Correlation between the simulated average growing-season stress factor and the simulated grain yield (the simulated stress-factor values were binned to the nearest 0.1 values for the stress factor; the black dot represent average simulated yield within each stress-factor bin, and the top and the bottom bars represent the 90th and 10th percentile values of the simulated grain yields within each bin; larger values of the stress factor indicate less constraints to plant growth due to N or water deficiency and smaller values of the stress factor indicate more constraints due to N or water deficiency; the blue line is linear regression model between the simulated stress factor and the grain yield and the grey shade indicate 95% confidence interval).

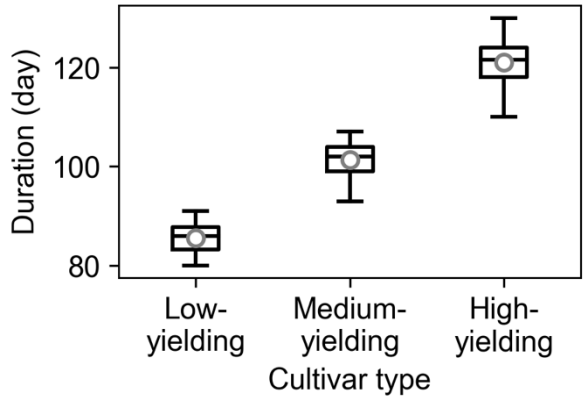

Figure 30 Distribution of the simulated maize yield of three maize cultivars in 1981-2010.

Table 6 Descriptions of the soil properties in the top 15 cm used in the yield forecasting algorithm

Site and soil types	Profile depth (cm)	Clay (%)	Silt (%)	Organic carbon (%)
Morogoro				
Poor	40	15	11.5	0.40
Medium	40	36.5	8.5	0.81
Fertile	60	36.5	8.5	0.86
Extremely fertile	100	36.5	8.5	0.95
Kagera				
Poor	40	20	9.5	0.1
Medium	40	32	11.5	0.58
Fertile	60	32	11.5	0.75
Extremely fertile	60	28.5	15.0	1.22
Tanga				
Poor	40	24.5	14.5	0.46
Medium	40	29.5	17.5	0.84
Fertile	60	29.5	17.5	1.11
Extremely fertile	60	31.5	21.0	1.52

The questionnaire consisted of a set of questions asking questions related to sampling location geographic information, and within-season agronomic information and climatic characteristics. The field questionnaire survey also included detailed instructions to establishing experimental plots. We included the key questions and instructions in the questionnaire below (Table 7).

Table 7 Key questions in the field survey questionnaire regarding the geographic agronomic and climatic information of sampling fields

Questions	Choices (when applicable)
Geographic and general information about the field	
Record the sampling field latitude and longitude.	
Record crops in the field.	
Record the clockwise and counter clock-wise field area using the GPS	

Table 7 (cont'd)

Questions related to management	
When was maize planted?	*Morogoro* Early February Mid February Late February Early March Mid March *Kagera and Tanga* Late August Early September Mid September Late September Early October
Which maize variety?	Short duration (3 months or less) Long duration (more than 3 months)
Has this field been irrigated?	Yes No
Select all of the fertilizers that have been applied to the field	SA CAN NPK Urea None
Has cow manure been applied to the field?	Yes No
Questions related to plant status	
Has maize experienced water stress?	Yes, major Yes, minor No
Has maize experienced nitrogen stress?	Yes, major Yes, minor No
Has the sampling field had any problems with diseases?	Yes, major Yes, minor No
Has the sampling field had any problems with insects?	Yes, major Yes, minor No
Has the sampling field had any problems with weeds?	Yes, major Yes, minor No
Questions related to in-season weather characteristics	
How was the weather before maize was planted?	Dry and hot Dry and cold Average rain and temperature Wet and cold Wet and hot

Instructions to selection of the experimental plot

Stand at the southwest corner of the sampling field. Record the coordinate of the southwest corner of the sampling field. Now, face north. Walk northward and count your steps until arriving at the northwest corner. This is side 1. Record the number of steps for side 1. Now walk eastward and count the number of steps. This is side 2. Record the number of steps for side 2. Now, return to the southwest corner of the field. Walk northward for <a random number less than the total steps in side 1> steps. Stop, and turn to your right. Enter the plot by walking for <a random number less than the total steps in side 2 minus 7> steps. Mark your position. This will be the southwest corner of the experimental plot. You will create a 6x6 m experimental plot. Divide the experimental plot into 3x3 m quadrants. You will select two plants from each quadrant, and fill in the following information related to the experimental plot (Table 8).

Table 8 Key questions in the field survey questionnaire for the experimental plot within the sampling field

Questions	Choices (when applicable)
Record the GPS coordinates of the southwest corner of the experimental plot.	
The questions below repeat for each quadrant	
What the total number of plants?	
Take a photo of the experimental plot.	
The questions below repeat for each of the two randomly sampled maize plants	
Has the plant tasseled?	Yes No
How many cobs are there?	
Are the cobs ready to be harvested?	Yes No
If the cobs are ready to be harvested, randomly choose two cobs and record the number of rows and the number of kernels per row.	
If the cobs on both plants are ready to be harvested, what is the total weight in grams of the cobs?	

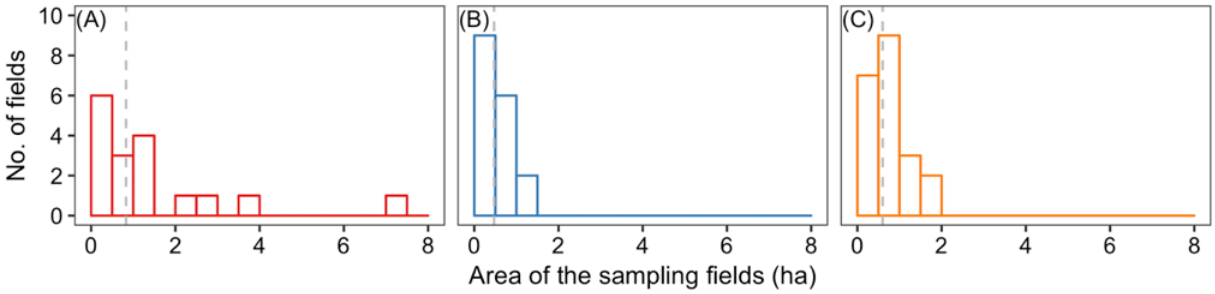

Figure 31 Distribution of sampling fields' size across (A) Morogoro, (B) Kagera and (C) Tanga (the size of 7 fields in Kagera were missing).

APPENDIX B: Chapter 3 Supplemental Tables and Figures

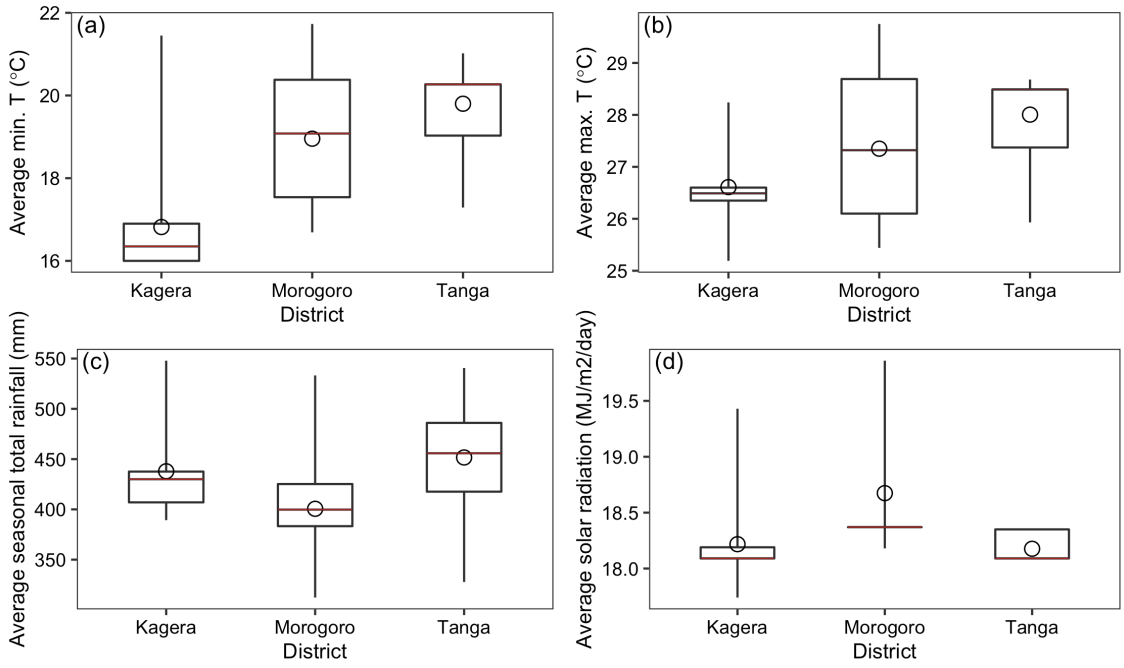

Figure 32 Distribution of average values of climatic variables in masika seasons (March-May) in 1990-2019 across 60 sites in the three districts: (a) minimum temperature, (b) maximum temperature, (c) seasonal total rainfall, and (d) daily solar radiation.

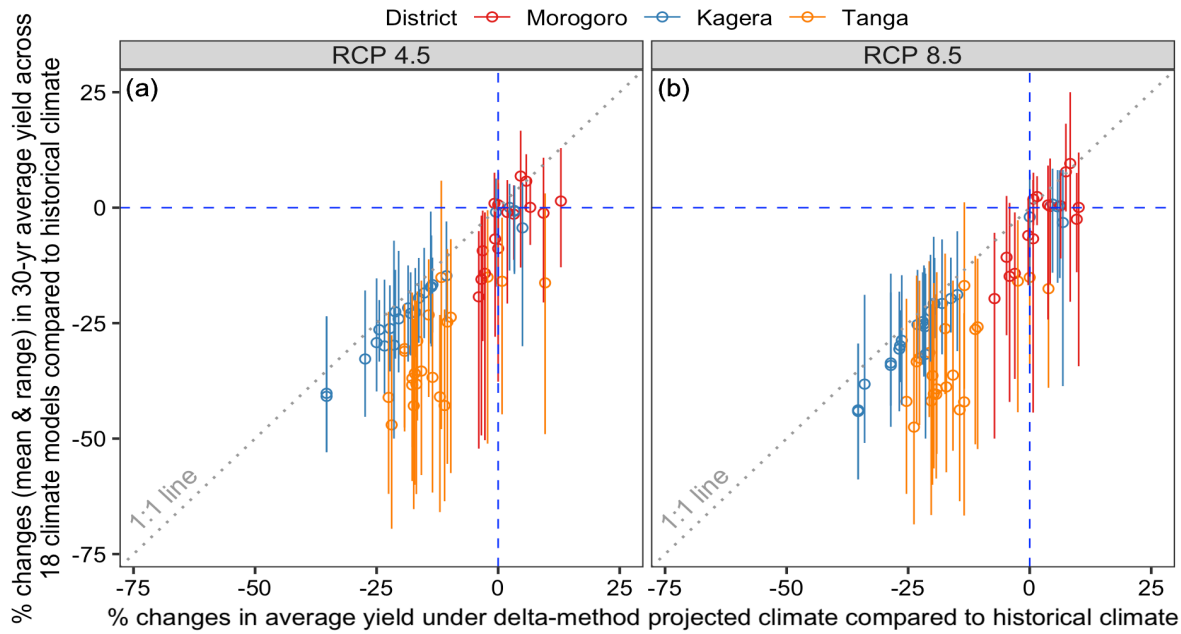

Figure 33 Comparisons of simulated changes in 30-year average yield in 2020-2049 under (a) RCP 4.5 and (b) RCP 8.5 compared to historical yield in 1990-2019 when using 18 climate models versus using one delta-method projected climate (the two ends of lines indicate minimum and maximum values, and the height of the circle on y-axis indicate the average value)

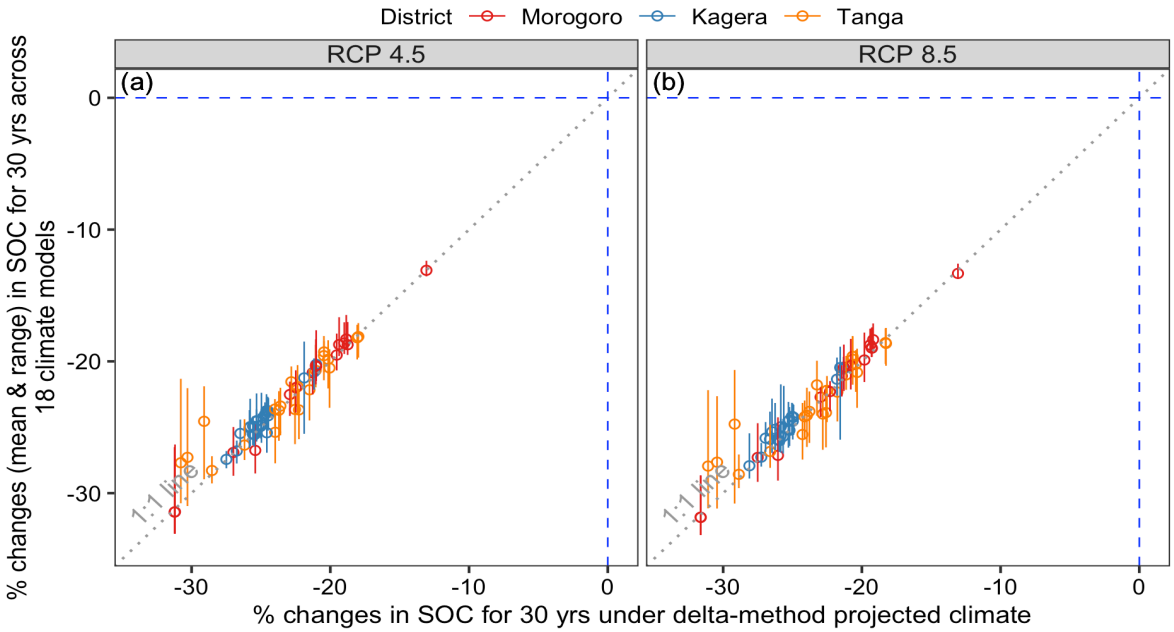

Figure 34 Comparisons of simulated SOC changes over 30 years under climate change when using 18 climate models versus using one delta-method projected climate (the two ends of lines indicate minimum and maximum values, and the height of the circle on y-axis indicate the average value)

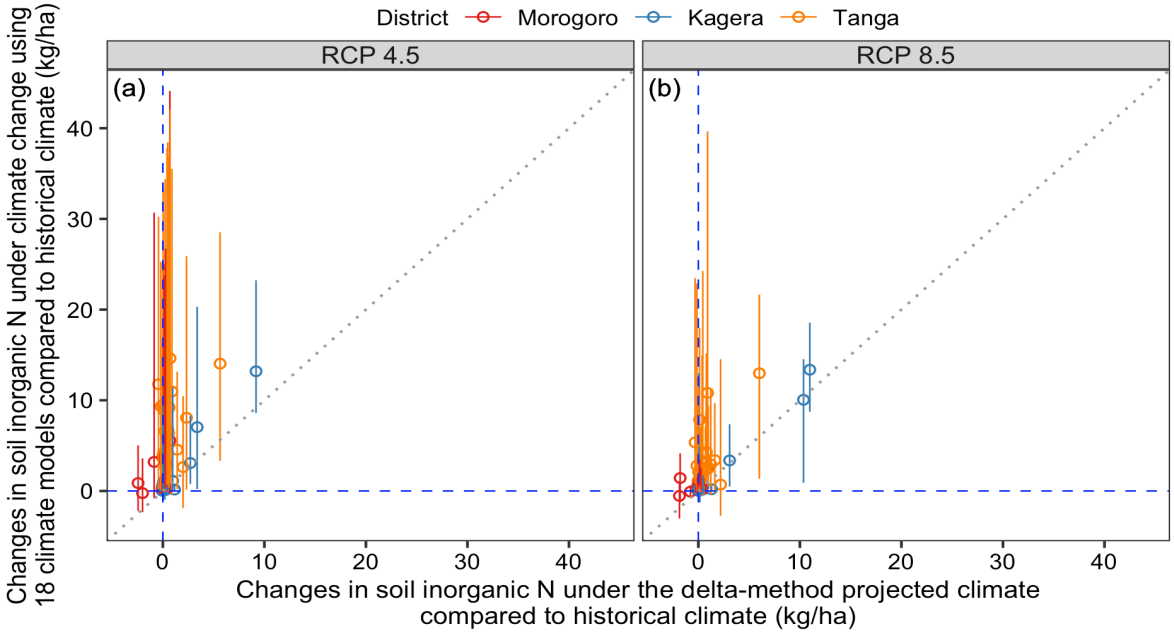

Figure 35 Comparisons of simulated changes in soil inorganic N under climate change compared to historical climate when using 18 climate models versus using one delta-method projected climate (the two ends of lines indicate minimum and maximum values, and the height of the circle on y-axis indicate the average value)

APPENDIX C: Chapter 4 Supplemental Tables and Figures

SALUS model did not count for biotic stress on crop growth. To capture the effect of weeds on yam growth, we applied reduction factors based on the amount of weed present in the yam field. Both tuber and aboveground biomass yield was reduced by 25% with 2.5 Mg/ha or more weeds. The biomass reduction factor was 20% with 2-2.4 Mg/ha weeds, 15% with 1.5-2 Mg/ha, and 10% with less than 1.5 Mg/ha weeds (Akobundu, 1981). Weed stress varied across treatments and years (Figure 36).

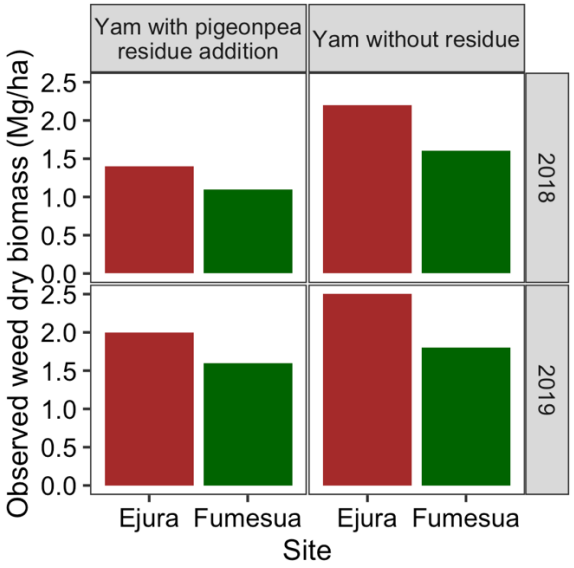

Figure 36 Observed weed biomass in two study sites under the two cropping systems (with and without pigeonpea residue addition) in 2018-2019

We compared available weather record of 20 weather stations to POWER weather data across Ghana in 2010-2014. The reported POWER annual precipitation matched with the recorded annual total precipitation by the weather stations in general (Figure 37a). We found drizzle events (<1mm daily precipitation) in POWER across Ghana. We also found the average minimum temperature was 1.3% higher in POWER dataset compared to the weather station record, and maximum temperature was 3% lower in the POWER dataset. We tested the effect of removing drizzle events and adjusting temperature on simulated yam yield on 20 POWER

climate grids where the weather stations were located. We used four climate scenarios, original POWER weather, drizzle events removed, temperature corrected using the percentage difference, and adjustment for both drizzle and temperature, to simulate 10-year continue rainfed unfertilized yam in 2005-2014. We did not find sizable impact of climate adjustment. The percentage difference in average 10-year simulated yam yield between adjusted POWER weather and original POWER weather was mostly (more than 97% cases) within 4% across the 20 stations (Figure 37b). Therefore, we used original POWER weather series.

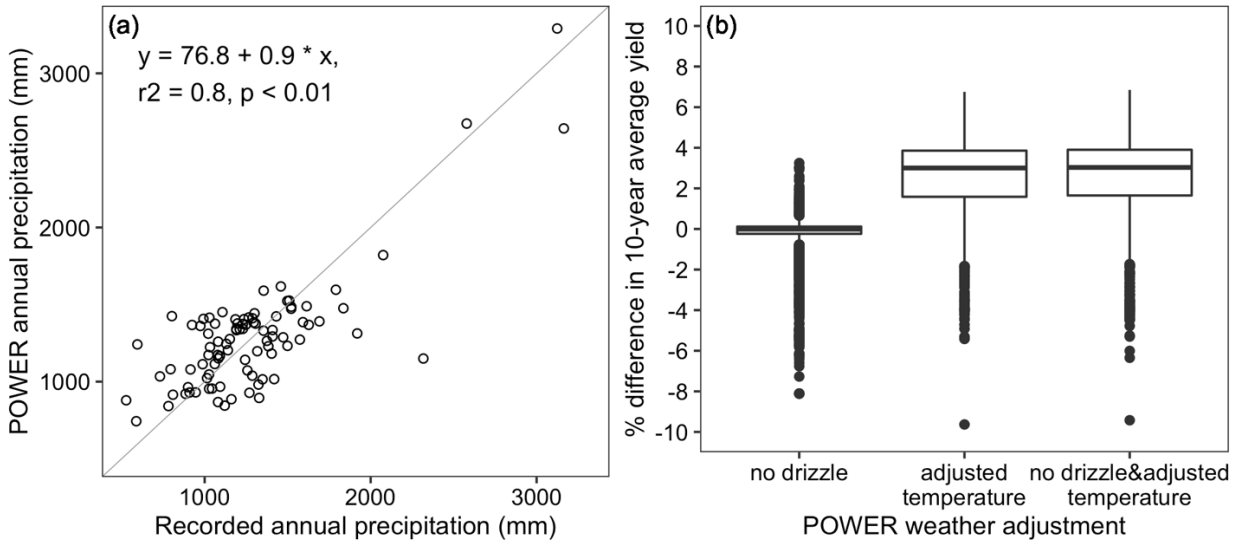

Figure 37 Percentage difference in 10-year average yam tuber yield between original POWER weather and adjusted POWER weather for drizzle, temperature or both.

We used coefficient of skewness to evaluate the normality of simulated 10 years of three growing-season stress levels --- drought stress, and N and P deficiency stress levels --- for each simulation unit. Growing-season stress level was the sum of daily stress level, which was computed as 1 – supply/demand, over a growing season. Skewness of 0 indicates symmetric (i.e. normal) distributions, positive skewness means right skewed and negative skewness means left skewed (Đorić et al., 2009; Fu et al., 2010). The drought stress level was not normally distributed (absolute skewness > 0.55) for 72% of the simulation units (Figure 38a) whereas the two nutrient

deficiency stress levels were close to normal distribution (absolute skewness <= 0.55) for about 65% of the simulation units (Figure 38b-c).

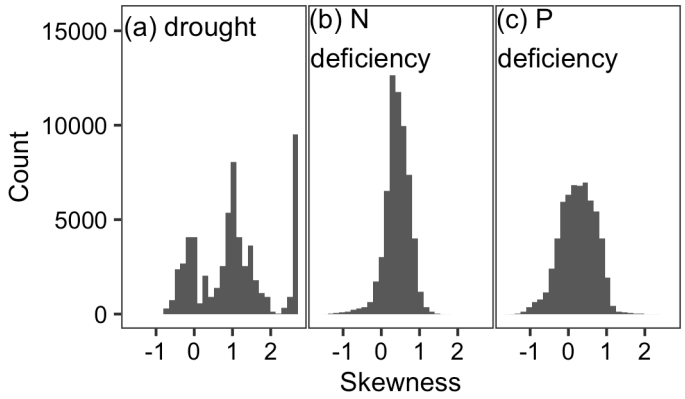

Figure 38 Distribution of skewness of 10 years of simulated stress levels for (a) drought, (b) N deficiency, and (c) P deficiency stress across each simulation unit (skewness value of 0 indicates normal distribution)

APPENDIX D: Chapter 5 Supplemental Tables and Figures

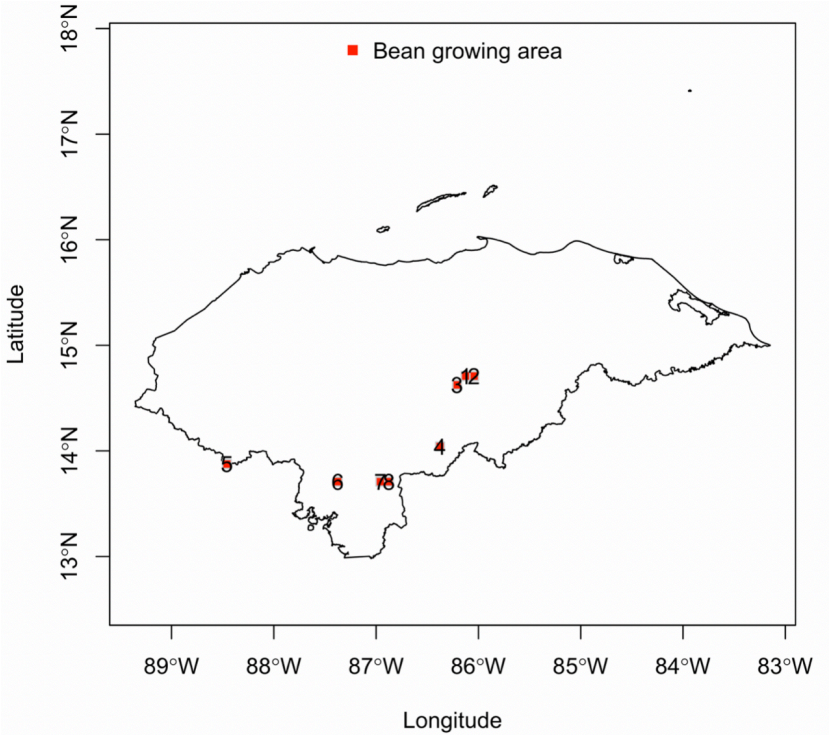

Figure 39 Spatial distribution of the eight 10x10km study regions where bean harvest area was over 700 ha in Honduras (IFPRI 2019; You et al., 2009)

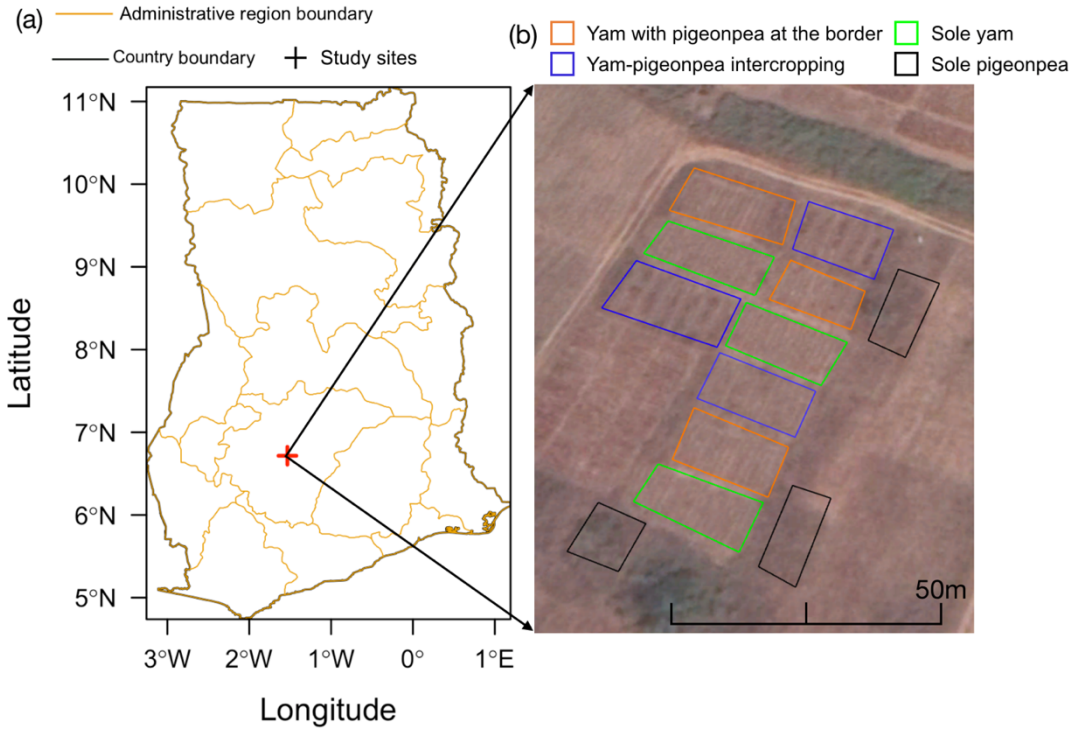

Figure 40 Study sites in Ghana (a) study location in Ghana, and (b) layout of the experimental plots with three pigeonpea cropping systems and three replicates per cropping system

The 8 bean growing areas in Honduras was characterized with distinct precipitation patterns. The area 1-4 received average annual total precipitation of about 1000mm and area 5-8 received average of 1200-1700mm in 2010-2018 (Figure 41). Beans are often planted in the secondary rain season in September in Honduras (Díaz-Ambrona et al., 2013).

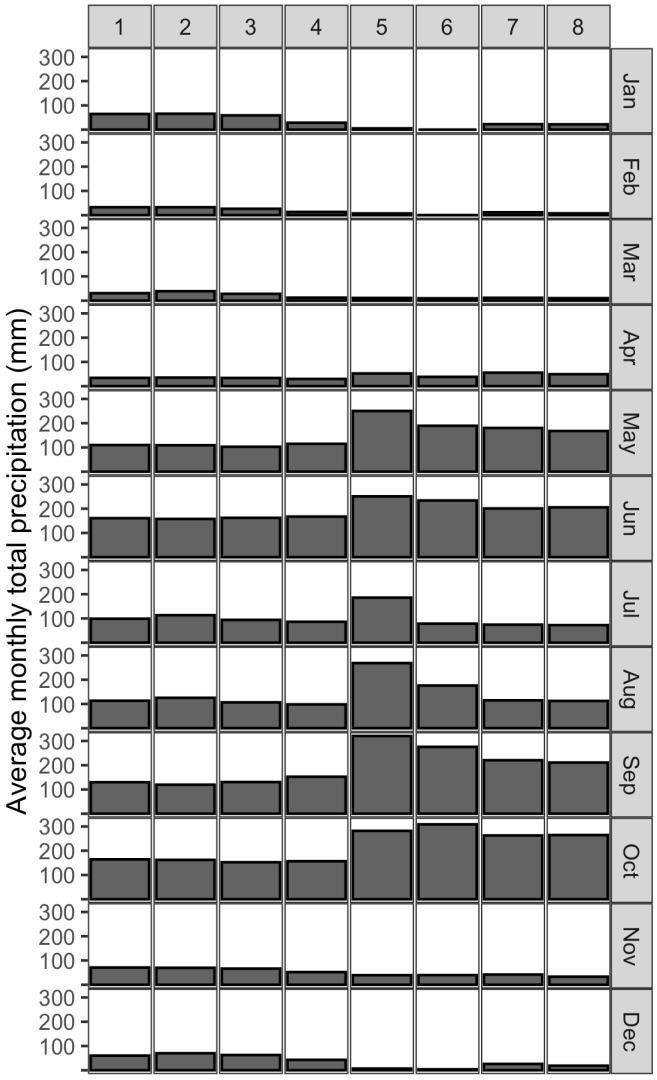

Figure 41 Average monthly total precipitation in 2011-2018 across the 8 bean growing areas (each column represents one bean growing area)

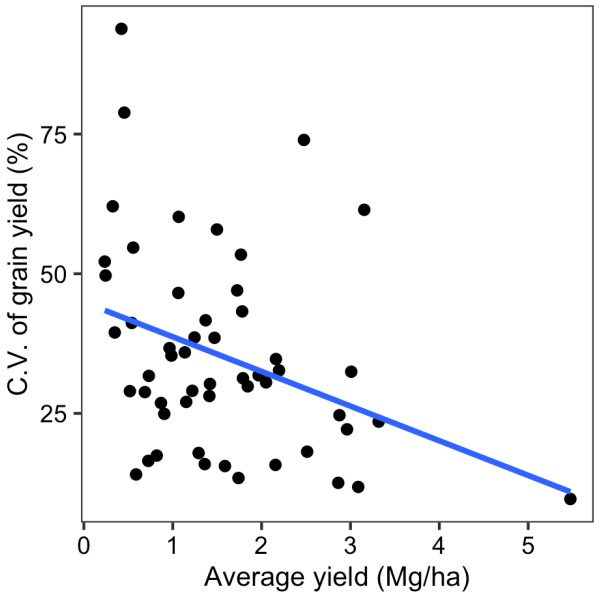

Figure 42 CV of grain yield versus average yield across study sites in Tanzania (blue line is fitted regression line using the general linear model)

REFERENCES

REFERENCES

Abate, T., Fisher, M., Abdoulaye, T., Kassie, G. T., Lunduka, R., Marenya, P., and Asnake, W. (2017). Characteristics of maize cultivars in Africa: How modern are they and how many do smallholder farmers grow? *Agriculture & Food Security* **6**, 30.

Abbam, T., Amoako Johnson, F., Dash, J., and Padmadas, S. (2018). Spatiotemporal Variations in Rainfall and Temperature in Ghana Over the Twentieth Century, 1900–2014. *Earth and Space Science* **5**, 120-132.

Abdoulaye, T., Alene, A., Rusike, J., and Akinola, A. A. (2014). "Strategic assessment of yam research priorities," Rep. No. 2309-6586, Lima Peru.

Abo-Shetaia, A. M., Ashoub, M. A., Ismail, M., & Al-Khaled, E. A. (2005). Estimation of some summer crops area and yield prediction using remote sensing techniques. *Annals of Agricultural Science (Cairo)* **50**, 481–498.

Abraham, M., and Pingali, P. (2020). Transforming Smallholder Agriculture to Achieve the SDGs. *In* "The Role of Smallholder Farms in Food and Nutrition Security" (S. Gomez y Paloma, L. Riesgo and K. Louhichi, eds.), pp. 173-209. Springer International Publishing, Cham.

Acheampong, P. P., Owusu Danquah, E., Dissanayake, H. G., Hayford, P., and Weebadde, C. (2019). A Socioeconomic Study of Transition Zone Yam Farmers Addressing Constraints and Exploring Opportunities for Integrating Pigeonpea into Yam Cropping Systems. *Sustainability* **11**, 717.

Adams, E. A., Kuusaana, E. D., Ahmed, A., and Campion, B. B. (2019). Land dispossessions and water appropriations: Political ecology of land and water grabs in Ghana. *Land Use Policy* **87**, 104068.

Adjei-Gyapong, T., and Asiamah, R. D. (2002). The interim Ghana soil classification system and its relation with the World Reference Base for Soil Resources. *Rapport sur les Ressources en Sols du Monde (FAO)*.

Adjei-Nsiah, S. (2012). Role of pigeonpea cultivation on soil fertility and farming system sustainability in Ghana. *International Journal of Agronomy* **2012**.

Agbede, T. M., Adekiya, A. O., and Ogeh, J. S. (2013). Effects of organic fertilizers on yam productivity and some soil properties of a nutrient-depleted tropical Alfisol. *Archives of Agronomy and Soil Science* **59**, 803-822.

Akobundu, I. O. (1981). Weed interference and control in white yam (Dioscorea rotundata Poir). *Weed Research* **21**, 267-272.

Albarenque, S. M., Basso, B., Caviglia, O. P., and Melchiori, R. J. M. (2016). Spatio-Temporal Nitrogen Fertilizer Response in Maize: Field Study and Modeling Approach. *Agronomy Journal* **108**, 2110-2122.

Amekudzi, L. K., Yamba, E. I., Preko, K., Asare, E. O., Aryee, J., Baidu, M., and Codjoe, S. N. A. (2015). Variabilities in rainfall onset, cessation and length of rainy season for the various agro-ecological zones of Ghana. *Climate* **3**, 416-434.

Anikwe, M. A. N. (2010). Carbon storage in soils of Southeastern Nigeria under different management practices. *Carbon balance and management* **5**, 5.

Arkin, G. F., Maas, S. J., and Richardson, C. W. (1980). Forecasting grain sorghum yields using simulated weather data and updating techniques. *Transactions of the ASAE* **23**, 676-0680.

Arndt, C., Farmer, W., Strzepek, K., and Thurlow, J. (2012). Climate change, agriculture and food security in Tanzania. *Review of Development Economics* **16**, 378-393.

Asseng, S., Cammarano, D., Basso, B., Chung, U., Alderman, P. D., Sonder, K., Reynolds, M., and Lobell, D. B. (2017). Hot spots of wheat yield decline with rising temperatures. *Global change biology* **23**, 2464-2472.

Asseng, S., Zhu, Y., Basso, B., Wilson, T., and Cammarano, D. (2014). Simulation modeling: applications in cropping systems. *In* "Encyclopedia of Agriculture and Food Systems" (N. V. Alfen, ed.), Vol. 5, pp. 102-112. Elsevier, San Diego.

Azzari, G., Jain, M., and Lobell, D. B. (2017). Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries. *Remote Sensing of Environment* **202**, 129-141.

Basso, B., Cammarano, D., Troccoli, A., Chen, D., and Ritchie, J. T. (2010). Long-term wheat response to nitrogen in a rainfed Mediterranean environment: Field data and simulation analysis. *European Journal of Agronomy* **33**, 132-138.

Basso, B., Dumont, B., Maestrini, B., Shcherbak, I., Robertson, G. P., Porter, J. R., Smith, P., Paustian, K., Grace, P. R., and Asseng, S. (2018a). Soil organic carbon and nitrogen feedbacks on crop yields under climate change. *Agricultural & Environmental Letters* **3**.

Basso, B., Dumont, B., Maestrini, B., Shcherbak, I., Robertson, G. P., Porter, J. R., Smith, P., Paustian, K., Grace, P. R., Asseng, S., Bassu, S., Biernath, C., Boote, K. J., Cammarano, D., De Sanctis, G., Durand, J. L., Ewert, F., Gayler, S., Hyndman, D. W., Kent, J., Martre, P., Nendel, C., Priesack, E., Ripoche, D., Ruane, A. C., Sharp, J., Thorburn, P. J., Hatfield, J. L., Jones, J. W., and Rosenzweig, C. (2018b). Soil Organic Carbon and Nitrogen Feedbacks on Crop Yields under Climate Change. *Agricultural & Environmental Letters* **3**, 180026.

Basso, B., Fiorentino, C., Cammarano, D., and Schulthess, U. (2016a). Variable rate nitrogen fertilizer response in wheat using remote sensing. *Precision agriculture* **17**, 168-182.

Basso, B., Gargiulo, O., Paustian, K., Robertson, G. P., Porter, C., Grace, P. R., and Jones, J. W. (2011). Procedures for initializing soil organic carbon pools in the DSSAT-CENTURY model for agricultural systems. *Soil Science Society of America Journal* **75**, 69-78.

Basso, B., Giola, P., Dumont, B., Migliorati, M. D. A., Cammarano, D., Pruneddu, G., and Giunta, F. (2016b). Tradeoffs between maize silage yield and nitrate leaching in a Mediterranean nitrate-vulnerable zone under current and projected climate scenarios. *PloS one* **11**, e0146360.

Basso, B., and Liu, L. (2019). Seasonal crop yield forecast: Methods, applications, and accuracies. *In* "Advances in Agronomy", Vol. 154, pp. 201-255. Academic Press.

Basso, B., Liu, L., and Ritchie, J. T. (2016c). A Comprehensive Review of the CERES-Wheat, -Maize and -Rice Models' Performances. *In* "Advances in Agronomy" (L. S. Donald, ed.), Vol. Volume 136, pp. 27-132. Academic Press.

Basso, B., and Ritchie, J. T. (2012). Assessing the Impact of Management Strategies on Water Use Efficiency Using Soil–Plant–Atmosphere Models. *Vadose Zone Journal* **11**, vzj2011.0173.

Basso, B., and Ritchie, J. T. (2015). Simulating crop growth and biogeochemical fluxes in response to land management using the SALUS model. *In* "The ecology of agricultural landscapes: long-term research on the path to sustainability" (S. K. Hamilton, J. E. Doll and G. P. Robertson, eds.), pp. 252-274. Oxford University Press, New York, NY USA.

Basso, B., Ritchie, J. T., Grace, P. R., and Sartori, L. (2006). Simulation of tillage systems impact on soil biophysical properties using the SALUS model. *Italian Journal of Agronomy* **1**, 677-688.

Bassu, S., Brisson, N., Durand, J.-L., Boote, K., and Lizaso, J. (2014). How do various maize crop models vary in their responses to climate change factors? *Global change biology* **20**, 2301-2320.

Bationo, A., Fening, J. O., and Kwaw, A. (2018). Assessment of Soil Fertility Status and Integrated Soil Fertility Management in Ghana. *In* "Improving the Profitability, Sustainability and Efficiency of Nutrients Through Site Specific Fertilizer Recommendations in West Africa Agro-Ecosystems: Volume 1" (A. Bationo, D. Ngaradoum, S. Youl, F. Lompo and J. O. Fening, eds.), pp. 93-138. Springer International Publishing, Cham.

Bationo, A., Kihara, J., Vanlauwe, B., Waswa, B., and Kimetu, J. (2007). Soil organic carbon dynamics, functions and management in West African agro-ecosystems. *Agricultural*

systems **94**, 13-25.

Bellon, M. R., Kotu, B. H., Azzarri, C., and Caracciolo, F. (2020). To diversify or not to diversify, that is the question. Pursuing agricultural development for smallholder farmers in marginal areas of Ghana. *World Development* **125**, 104682.

Berre, D., Corbeels, M., Rusinamhodzi, L., Mutenje, M., Thierfelder, C., and Lopez-Ridaura, S. (2017). Thinking beyond agronomic yield gap: Smallholder farm efficiency under contrasted livelihood strategies in Malawi. *Field crops research* **214**, 113-122.

Brown, B., Nuberg, I., and Llewellyn, R. (2019). Pathways to intensify the utilization of conservation agriculture by African smallholder farmers. *Renewable Agriculture and Food Systems* **34**, 558-570.

Bruun, T. B., Elberling, B., de Neergaard, A., and Magid, J. (2015). Organic carbon dynamics in different soil types after conversion of forest to agriculture. *Land Degradation & Development* **26**, 272-283.

Burke, M., and Lobell, D. B. (2017). Satellite-based assessment of yield variation and its determinants in smallholder African systems. *Proceedings of the National Academy of Sciences* **114**, 2189-2194.

Cannon, A. J. (2018). Multivariate Bias Correction of Climate Model Outputs.

Cannon, A. J., Sobie, S. R., and Murdock, T. Q. (2015). Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate* **28**, 6938-6959.

Carsky, R. J., Asiedu, R., and Cornet, D. (2010). Review of soil fertility management for yam-based systems in West Africa. *African Journal of Root and Tuber Crops* **8**, 1-17.

Choularton, R. J., and Krishnamurthy, P. K. (2019). How accurate is food security early warning? Evaluation of FEWS NET accuracy in Ethiopia. *Food Security* **11**, 333-344.

Cillis, D., Maestrini, B., Pezzuolo, A., Marinello, F., and Sartori, L. (2018). Modeling soil organic carbon and carbon dioxide emissions in different tillage systems supported by precision agriculture technologies under current climatic conditions. *Soil and Tillage Research* **183**, 51-59.

Coughlan de Perez, E., van Aalst, M., Choularton, R., van den Hurk, B., Mason, S., Nissan, H., and Schwager, S. (2019). From rain to famine: assessing the utility of rainfall observations and seasonal forecasts to anticipate food insecurity in East Africa. *Food Security* **11**, 57-68.

Danquah, E. O., Ennin, S. A., Frimpong, F., Akom, M., and Acheampong, P. P. (2018). Improved Agronomic Practices for Sustainable Yam Production: The on Farm

Experience. *Agricultural and Food Science Journal of Ghana* **11**, 904-908.

Daroub, S. H., Gerakis, A., Ritchie, J. T., Friesen, D. K., and Ryan, J. (2003). Development of a soil-plant phosphorus simulation model for calcareous and weathered tropical soils. *Agricultural Systems* **76**, 1157-1181.

Delincé, J. (2017). "Recent practices and advances for AMIS crop yield forecasting at farm/parcel level: A review.," Rome, Italy.

DelSole, T., Nattala, J., and Tippett, M. K. (2014). Skill improvement from increased ensemble size and model diversity. *Geophysical Research Letters* **41**, 7331-7342.

Diao, X. (2016). Economywide Impact of Maize Export Bans on Agricultural Growth and Household Welfare in Tanzania: A Dynamic Computable General Equilibrium Model Analysis. *Development policy review* **34**, 101-134.

Díaz-Ambrona, C. G. H., Gigena, R., and Mendoza, C. O. (2013). Climate change impacts on maize and dry bean yields of smallholder farmers in Honduras. *Revista iberoamericana de estudios de desarrollo= Iberoamerican journal of development studies* **2**, 4-22.

Diby, L. N., Hgaza, V. K., Tié, T. B., Assa, A., Carsky, R., Girardin, O., Sangakkara, U. R., and Frossard, E. (2011). How does soil fertility affect yam growth? *Acta Agriculturae Scandinavica, Section B-Soil & Plant Science* **61**, 448-457.

Đorić, D., Nikolić-Đorić, E., Jevremović, V., and Mališić, J. (2009). On measuring skewness and kurtosis. *Quality and Quantity* **43**, 481-493.

Duan, T., Chapman, S. C., Guo, Y., and Zheng, B. (2017). Dynamic monitoring of NDVI in wheat agronomy and breeding trials using an unmanned aerial vehicle. *Field Crops Research* **210**, 71-80.

Duncan, J., Dash, J., and Atkinson, P. M. (2015). The potential of satellite-observed crop phenology to enhance yield gap assessments in smallholder landscapes. *Frontiers in Environmental Science* **3**, 56.

Dzotsi, K. A., Basso, B., and Jones, J. W. (2013). Development, uncertainty and sensitivity analysis of the simple SALUS crop model in DSSAT. *Ecological Modelling* **260**, 62-76.

Ehsan, M. A., Tippett, M. K., Almazroui, M., Ismail, M., Yousef, A., Kucharski, F., Omar, M., Hussein, M., and Alkhalaf, A. A. (2017). Skill and predictability in multimodel ensemble forecasts for Northern Hemisphere regions with dominant winter precipitation. *Climate Dynamics* **48**, 3309-3324.

Ennin, S. A., Issaka, R. N., Acheampong, P. P., Numafo, M., and Owusu Danquah, E. (2014). Mechanization, fertilization and staking options for environmentally sound yam production. *African Journal Agricultural Research* **9**, 2222-2230.

Falconnier, G.N., Corbeels, M., Boote, K.J., Affholder F., Adam M., MacCarthy D.S., Ruane A.C., Nendel C., Whitbread A.M., Justes E., Ahuja L.R., Akinseye F.M., Alou I.N., Amouzou K.A., Anapalli S.S., Baron C., Basso B., Baudron F., Bertuzzi P., …, Webber H. (in press) Modelling climate change impacts on maize yields under low nitrogen input conditions in sub-Saharan Africa, *Climate Change Biology*

FAO (1996). "Food, agriculture and food security: developments since the World Food Conference and prospects," Rome, Italy.

FAO (2005). "Fertilizer use by crop in Ghana," Rome, Italy.

Folberth, C., Skalský, R., Moltchanova, E., Balkovič, J., Azevedo, L. B., Obersteiner, M., and van der Velde, M. (2016). Uncertainty in soil data can outweigh climate impact signals in global crop yield simulations. *Nature Communications* **7**, 11872.

Fritz, S., See, L., McCallum, I., You, L., Bun, A., Moltchanova, E., Duerauer, M., Albrecht, F., Schill, C., Perger, C., Havlik, P., Mosnier, A., Thornton, P., Wood-Sichra, U., Herrero, M., Becker-Reshef, I., Justice, C., Hansen, M., Gong, P., Abdel Aziz, S., Cipriani, A., Cumani, R., Cecchi, G., Conchedda, G., Ferreira, S., Gomez, A., Haffani, M., Kayitakire, F., Malanding, J., Mueller, R., Newby, T., Nonguierma, A., Olusegun, A., Ortner, S., Rajak, D. R., Rocha, J., Schepaschenko, D., Schepaschenko, M., Terekhov, A., Tiangwa, A., Vancutsem, C., Vintrou, E., Wenbin, W., van der Velde, M., Dunwoody, A., Kraxner, F., and Obersteiner, M. (2015). Mapping global cropland and field size. *Global Change Biology* **21**, 1980-1992.

Frossard, E., Aighewi, B. A., Aké, S., Barjolle, D., Baumann, P., Bernet, T., Dao, D., Diby, L. N., Floquet, A., Hgaza, V. K., Ilboudo, L. J., Kiba, D. I., Mongbo, R. L., Nacro, H. B., Nicolay, G. L., Oka, E., Ouattara, Y. F., Pouya, N., Senanayake, R. L., Six, J., and Traoré, O. I. (2017). The Challenge of Improving Soil Fertility in Yam Cropping Systems of West Africa. *Frontiers in plant science* **8**, 1953-1953.

Fu, W., Tunney, H., and Zhang, C. (2010). Spatial variation of soil nutrients in a dairy farm and its implications for site-specific fertilizer application. *Soil and Tillage Research* **106**, 185-193.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., and Hoell, A. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific data* **2**, 150066.

Gennari, P., and Fonteneau, F. (2016). "Crop Yield Forecasting: Methodological and Institutional Aspects," Rome, Italy.

Gilardelli, C., Stella, T., Confalonieri, R., Ranghetti, L., Campos-Taberner, M., García-Haro, F. J., and Boschetti, M. (2019). Downscaling rice yield simulation at sub-field scale using remotely sensed LAI data. *European Journal of Agronomy* **103**, 108-116.

Giola, P., Basso, B., Pruneddu, G., Giunta, F., and Jones, J. W. (2012). Impact of manure and slurry applications on soil nitrate in a maize–triticale rotation: Field study and long term simulation analysis. *European Journal of Agronomy* **38**, 43-53.

Giorgi, F. (2019). Thirty Years of Regional Climate Modeling: Where Are We and Where Are We Going next? *Journal of Geophysical Research: Atmospheres* **124**, 5696-5723.

Giorgi, F., Jones, C., and Asrar, G. R. (2009). Addressing climate information needs at the regional level: the CORDEX framework. *World Meteorological Organization (WMO) Bulletin* **58**, 175.

Godwin, R. J., and Miller, P. C. H. (2003). A review of the technologies for mapping within-field variability. *Biosystems engineering* **84**, 393-407.

Hansen, M. C., Egorov, A., Potapov, P. V., Stehman, S. V., Tyukavina, A., Turubanova, S. A., Roy, D. P., Goetz, S. J., Loveland, T. R., and Ju, J. (2014). Monitoring conterminous United States (CONUS) land cover change with web-enabled Landsat data (WELD). *Remote sensing of Environment* **140**, 466-484.

Hemer, M. A., Fan, Y., Mori, N., Semedo, A., and Wang, X. L. (2013). Projected changes in wave climate from a multi-model ensemble. *Nature Climate Change* **3**, 471-476.

Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J. G. B., and Walsh, M. G. (2014). SoilGrids1km—global soil information based on automated mapping. *PloS one* **9**, e114788.

Hengl, T., Leenaars, J. G. B., Shepherd, K. D., Walsh, M. G., Heuvelink, G. B. M., Mamo, T., Tilahun, H., Berkhout, E., Cooper, M., Fegraus, E., Wheeler, I., and Kwabena, N. A. (2017). Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. *Nutrient Cycling in Agroecosystems* **109**, 77-102.

IFPRI (2019). Global Spatially-Disaggregated Crop Production Statistics Data for 2010 Version 1.1. Harvard Dataverse, V3.

Iizumi, T., Kotoku, M., Kim, W., West, P. C., Gerber, J. S., and Brown, M. E. (2018). Uncertainties of potentials and recent changes in global yields of major crops resulting from census-and satellite-based yield datasets at multiple resolutions. *PloS one* **13**, e0203809.

Jain, M. (2020). The Benefits and Pitfalls of Using Satellite Data for Causal Inference. *Review of Environmental Economics and Policy* **14**, 157-169.

Jayne, T. S., and Rashid, S. (2010). "The value of accurate crop production forecasts." Michigan State University, East Lansing, MI, USA.

Jégo, G., Pattey, E., and Liu, J. (2012). Using Leaf Area Index, retrieved from optical imagery, in the STICS crop model for predicting yield and biomass of field crops. *Field Crops Research* **131**, 63-74.

Jin, Z. (2019). Smallholder maize area and yield mapping at national scales with Google Earth Engine. *Remote sensing of environment* **228**, 115-128.

Johnson, D. M. (2014). An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. *Remote Sensing of Environment* **141**, 116-128.

Jones, J. W., Antle, J. M., Basso, B., Boote, K. J., Conant, R. T., Foster, I., Godfray, H. C. J., Herrero, M., Howitt, R. E., Janssen, S., Keating, B. A., Munoz-Carpena, R., Porter, C. H., Rosenzweig, C., and Wheeler, T. R. (2017). Brief history of agricultural systems modeling. *Agricultural Systems* **155**, 240-254.

Kadaja, J., Saue, T., and Vii, P. (2009). PROBABILISTIC YIELD FORECAST BASED ON A PRODUCTION PROCESS MODEL. *In* "Computer and Computing Technologies in Agriculture Ii, Vol 1" (D. Li and C. Zhao, eds.), Vol. 293, pp. 487-494.

Kahsay, G. A., and Hansen, L. G. (2016). The effect of climate change and adaptation policy on agricultural production in Eastern Africa. *Ecological Economics* **121**, 54-64.

Katz, R. W. (1977). Assessing the impact of climatic change on food production. *Climatic Change* **1**, 85-96.

Kayode, R. M. O., Buhari, O. J., Otutu, L. O., Ajibola, T. B., Oyeyinka, S. A., Opaleke, D. O., and Akeem, S. A. (2017). Physicochemical properties of processed aerial yam (Dioscorea bulbifera) and sensory properties of paste (amala) prepared with cassava flour. *Journal of Agricultural Sciences–Sri Lanka* **12**, 84-94.

Kesh, H., Yadav, A. S., Sarial, A. K., Khajuria, S., and Jain, B. T. (2017). Variability for Nitrogen and Phosphorus Content in Pigeon Pea (Cajanus cajan L.) in Response to Rhizobium and Piriformospora indica. *Research Journal of Agricultural Sciences* **8**, 203-206.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift* **15**, 259-263.

Kuri, F., Murwira, A., Murwira, K. S., and Masocha, M. (2014). Predicting maize yield in Zimbabwe using dry dekads derived from remotely sensed Vegetation Condition Index. *International Journal of Applied Earth Observation and Geoinformation* **33**, 39-46.

Lambert, M.-J., Traoré, P. C. S., Blaes, X., Baret, P., and Defourny, P. (2018). Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sensing of Environment* **216**, 647-657.

Law-Ogbomo, K. E., and Egharevba, R. K. A. (2009). COMPARISON OF YIELD PERFORMANCE AND PROFITABILITY OF Dioscorea rotundata UNDER DIFFERENT NPK FERTILIZER REGIME IN HUMID FOREST ULTISOLS. *Tropical and Subtropical Agroecosystems* **10**, 451-456.

Law-Ogbomo, K. E., and Remison, S. U. (2008). Growth and yield of white guinea yam (Dioscorea rotundata Poir.) influenced by NPK fertilization on a forest site in Nigeria. *Journal of Tropical Agriculture* **46**, 21-24.

Law-Ogbomo, K. E., and Remison, S. U. (2009). Yield and distribution/uptake of nutrients of Dioscorea rotundata influenced by NPK fertilizer application. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* **37**, 165-170.

Lesk, C., Rowhani, P., and Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature* **529**, 84-87.

Liaghat, S., and Balasundram, S. K. (2010). A review: The role of remote sensing in precision agriculture. *American journal of agricultural and biological sciences* **5**, 50-55.

Liu, L., and Basso, B. (2017a). Spatial evaluation of maize yield in Malawi. *Agricultural Systems* **157**, 185-192.

Liu, L., and Basso, B. (2017b). Spatial evaluation of switchgrass productivity under historical and future climate scenarios in Michigan. *GCB Bioenergy* **9**, 1320-1332.

Liu, L., and Basso, B. (2020a). Impacts of climate variability and adaptation strategies on crop yields and soil organic carbon in the US Midwest. *PloS one* **15**, e0225433.

Liu, L., and Basso, B. (2020b). Linking field survey with crop modeling to forecast maize yield in smallholder farmers' fields in Tanzania. *Food Security*, 1-12.

Lobell, D. B., Thau, D., Seifert, C., Engle, E., and Little, B. (2015). A scalable satellite-based crop yield mapper. *Remote Sensing of Environment* **164**, 324-333.

Lourens, U. W., and De Jager, J. M. (1997). A computerized crop-specific drought monitoring system: Design concepts and initial testing. *Agricultural Systems* **53**, 303-315.

Lowder, S. K., Skoet, J., and Raney, T. (2016). The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Development* **87**, 16-29.

Luhunga, P., Chang'a, L., and Djolov, G. (2017). Assessment of the impacts of climate change on maize production in the Wami Ruvu basin of Tanzania. *Journal of Water and Climate Change* **8**, 142-164.

Luhunga, P. M. (2017). Assessment of the Impacts of Climate Change on Maize Production in the Southern and Western Highlands Sub-agro Ecological Zones of Tanzania. *Frontiers*

in Environmental Science **5**, 51.

Luhunga, P. M., Botai, J. O., and Kahimba, F. (2016). Evaluation of the performance of CORDEX regional climate models in simulating present climate conditions of Tanzania. *Journal of Southern Hemisphere Earth Systems Science* **66**, 32-54.

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S., and Schimel, D. S. (2011). Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications* **21**, 1429-1442.

Maestrini, B., and Basso, B. (2018a). Drivers of within-field spatial and temporal variability of crop yield across the US Midwest. *Scientific reports* **8**, 1-9.

Maestrini, B., and Basso, B. (2018b). Predicting spatial patterns of within-field crop yield variability. *Field Crops Research* **219**, 106-112.

Maliki, R., Toukourou, M., Sinsin, B., and Vernier, P. (2012). Productivity of yam-based systems with herbaceous legumes and short fallows in the Guinea-Sudan transition zone of Benin. *Nutrient cycling in Agroecosystems* **92**, 9-19.

Manatsa, D., Nyakudya, I. W., Mukwada, G., and Matsikwa, H. (2011). Maize yield forecasting for Zimbabwe farming sectors using satellite rainfall estimates. *Natural hazards* **59**, 447-463.

Marcos, J., Cornet, D., Bussière, F., and Sierra, J. (2011). Water yam (Dioscorea alata L.) growth and yield as affected by the planting date: Experiment and modelling. *European Journal of Agronomy* **34**, 247-256.

Mason-D'Croz, D., Sulser, T. B., Wiebe, K., Rosegrant, M. W., Lowder, S. K., Nin-Pratt, A., Willenbockel, D., Robinson, S., Zhu, T., Cenacchi, N., Dunston, S., and Robertson, R. D. (2019). Agricultural investments and hunger in Africa modeling potential contributions to SDG2 – Zero Hunger. *World Development* **116**, 38-53.

Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., and Pidcock, R. (2018). Global warming of 1.5 C. *An IPCC Special Report on the impacts of global warming of* **1**.

Michler, J. D., Baylis, K., Arends-Kuenning, M., and Mazvimavi, K. (2019). Conservation agriculture and climate resilience. *Journal of Environmental Economics and Management* **93**, 148-169.

Mishra, A., Hansen, J. W., Dingkuhn, M., Baron, C., Traore, S. B., Ndiaye, O., and Ward, M. N. (2008). Sorghum yield prediction from seasonal rainfall forecasts in Burkina Faso. *Agricultural and Forest Meteorology* **148**, 1798-1814.

Mkhabela, M. S., Mkhabela, M. S., and Mashinini, N. N. (2005). Early maize yield

forecasting in the four agro-ecological regions of Swaziland using NDVI data derived from NOAA's-AVHRR. *Agricultural and Forest Meteorology* **129**, 1-9.

MoFA (2017). "Agriculture in Ghana: Facts and Figures." Statistics Research and Information Directorate (SRID) of the Ministry of Food and Agriculture, Accra, Ghana.

Moore, N., Alagarswamy, G., Pijanowski, B., Thornton, P., Lofgren, B., Olson, J., Andresen, J., Yanda, P., and Qi, J. (2012). East African food security as influenced by future climate change and land use change at local to regional scales. *Climatic change* **110**, 823-844.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE* **50**, 885-900.

Mourice, S. K., Mbungu, W., and Tumbo, S. D. (2017). Quantification of Climate Change and Variability Impacts on Maize Production at Farm Level in the Wami River Sub-Basin, Tanzania. *In* "Quantification of Climate Variability, Adaptation and Mitigation for Agricultural Sustainability" (M. Ahmed and C. O. Stockle, eds.), pp. 323-351. Springer International Publishing, Cham.

Msongaleli, B. M., Rwehumbiza, F., Tumbo, S. D., and Kihupi, N. (2015). Impacts of climate variability and change on rainfed sorghum and maize: implications for food security policy in Tanzania.

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering* **114**, 358-371.

Naab, J. B., Mahama, G. Y., Koo, J., Jones, J. W., and Boote, K. J. (2015). Nitrogen and phosphorus fertilization with crop residue retention enhances crop productivity, soil organic carbon, and total soil nitrogen concentrations in sandy-loam soils in Ghana. *Nutrient cycling in agroecosystems* **102**, 33-43.

Nakamura, S., Hayashi, K., Omae, H., Ramadjita, T., Dougbedji, F., Shinjo, H., Saidou, A. K., and Tobita, S. (2011). Validation of soil organic carbon dynamics model in the semi-arid tropics in Niger, West Africa. *Nutrient cycling in agroecosystems* **89**, 375-385.

Nandram, B., Berg, E., and Barboza, W. (2014). A hierarchical Bayesian model for forecasting state-level corn yield. *Environmental and ecological statistics* **21**, 507-530.

Nhemachena, C., Matchaya, G., Nhemachena, C. R., Karuaihe, S., Muchara, B., and Nhlengethwa, S. (2018). Measuring baseline agriculture-related sustainable development goals index for southern Africa. *Sustainability* **10**, 849.

Nkonya, E. (1998). "Adoption of maize production technologies in Northern Tanzania," CIMMYT.

Olatoye, K. K., and Arueya, G. L. (2019). Nutrient and phytochemical composition of flour made from selected cultivars of Aerial yam (Dioscorea bulbifera) in Nigeria. *Journal of Food Composition and Analysis* **79**, 23-27.

Ouédraogo, E., Mando, A., Brussaard, L., and Stroosnijder, L. (2007). Tillage and fertility management effects on soil organic matter and sorghum yield in semi-arid West Africa. *Soil and Tillage Research* **94**, 64-74.

Owusu, S., Yigini, Y., Olmedo, G. F., and Omuto, C. T. (2020). Spatial prediction of soil organic carbon stocks in Ghana using legacy data. *Geoderma* **360**, 114008.

Paavola, J. (2008). Livelihoods, vulnerability and adaptation to climate change in Morogoro, Tanzania. *Environmental Science & Policy* **11**, 642-654.

Pease, J. W., Wade, E. W., Skees, J. S., and Shrestha, C. M. (1993). Comparisons between subjective and statistical forecasts of crop yields. *Review of agricultural economics* **15**, 339-350.

Persello, C., Tolpekin, V. A., Bergado, J. R., and de By, R. A. (2019). Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. *Remote sensing of environment* **231**, 111253.

Pezzuolo, A., Dumont, B., Sartori, L., Marinello, F., De Antoni Migliorati, M., and Basso, B. (2017). Evaluating the impact of soil conservation measures on soil organic carbon at the farm scale. *Computers and Electronics in Agriculture* **135**, 175-182.

Pfahl, S., O'Gorman, P. A., and Fischer, E. M. (2017). Understanding the regional pattern of projected future changes in extreme precipitation. *Nature Climate Change* **7**, 423-427.

Phiri, A. T., Njoloma, J. P., Kanyama-Phiri, G. Y., Snapp, S., and Lowole, M. W. (2010). Maize yield response to the combined application of Tundulu rock phosphate and Pigeon Pea residues in Kasungu, Central Malawi. *African Journal of Agricultural Research* **5**, 1235-1242.

Raun, W. R., Solie, J. B., Stone, M. L., Martin, K. L., Freeman, K. W., Mullen, R. W., Zhang, H., Schepers, J. S., and Johnson, G. V. (2005). Optical sensor-based algorithm for crop nitrogen fertilization. *Communications in Soil Science and Plant Analysis* **36**, 2759-2781.

Raymundo, R., Asseng, S., Cammarano, D., and Quiroz, R. (2014). Potato, sweet potato, and yam models for climate change: A review. *Field Crops Research* **166**, 173-185.

Reynolds, C. A., Yitayew, M., Slack, D. C., Hutchinson, C. F., Huete, A., and Petersen, M. S. (2000). Estimating crop yields and production by integrating the FAO Crop Specific Water Balance model with real-time satellite data and ground-based ancillary data. *International Journal of Remote Sensing* **21**, 3487-3508.

Rhebergen, T., Fairhurst, T., Zingore, S., Fisher, M., Oberthur, T., and Whitbread, A. (2016). Climate, soil and land-use based land suitability evaluation for oil palm production in Ghana. *European Journal of Agronomy* **81**, 1-14.

Ritchie, J. T. (1972). Model for predicting evaporation from a row crop with incomplete cover. *Water Resources Research* **8**, 1204-1213.

Ritchie, J. T., and Basso, B. (2008). Water use efficiency is not constant when crop water supply is adequate or fixed: The role of agronomic management. *European Journal of Agronomy* **28**, 273-281.

Ritchie, J. T., Porter, C. H., Judge, J., Jones, J. W., and Suleiman, A. A. (2009). Extension of an existing model for soil water evaporation and redistribution under high water content conditions. *Soil Science Society of America Journal* **73**, 792-801.

Robertson, M. J., Lyle, G., and Bowden, J. W. (2008). Within-field variability of wheat yield and economic implications for spatially variable nutrient management. *Field Crops Research* **105**, 211-220.

Rojas, O. (2007). Operational maize yield model development and validation based on remote sensing and agro-meteorological data in Kenya. *International Journal of Remote Sensing* **28**, 3775-3793.

Rowhani, P., Lobell, D. B., Linderman, M., and Ramankutty, N. (2011). Climate variability and crop production in Tanzania. *Agricultural and Forest Meteorology* **151**, 449-460.

Ruane, A. C., Goldberg, R., and Chryssanthacopoulos, J. (2015). Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. *Agricultural and Forest Meteorology* **200**, 233-248.

Schauberger, B., Gornott, C., and Wechsung, F. (2017). Global evaluation of a semi-empirical model for yield anomalies and application to within-season yield forecasting. *Global Change Biology* **23**, 4750-4764.

Schut, A. G. T., Traore, P. C. S., Blaes, X., and de By, R. A. (2018). Assessing yield and fertilizer response in heterogeneous smallholder fields with UAVs and satellites. *Field Crops Research* **221**, 98-107.

Serdeczny, O., Adams, S., Baarsch, F., Coumou, D., Robinson, A., Hare, W., Schaeffer, M., Perrette, M., and Reinhardt, J. (2017). Climate change impacts in Sub-Saharan Africa: from physical changes to their social repercussions. *Regional Environmental Change* **17**, 1585-1600.

Sheffield, J., Wood, E. F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., and Demuth, S. (2014). A drought monitoring and forecasting system for sub-Sahara African water resources and food security. *Bulletin of the American*

Meteorological Society **95**, 861-882.

Sibanda, M., and Murwira, A. (2012). The use of multi-temporal MODIS images with ground data to distinguish cotton from maize and sorghum fields in smallholder agricultural landscapes of Southern Africa. *International Journal of Remote Sensing* **33**, 4841-4855.

Sillmann, J., Thorarinsdottir, T., Keenlyside, N., Schaller, N., Alexander, L. V., Hegerl, G., Seneviratne, S. I., Vautard, R., Zhang, X., and Zwiers, F. W. (2017). Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities. *Weather and climate extremes* **18**, 65-74.

Singh, C., Daron, J., Bazaz, A., Ziervogel, G., Spear, D., Krishnaswamy, J., Zaroug, M., and Kituyi, E. (2018). The utility of weather and climate information for adaptation decision-making: current uses and future prospects in Africa and India. *Climate and Development* **10**, 389-405.

Sitko, N. J., Chisanga, B., Tschirley, D., and Jayne, T. S. (2018). An evolution in the middle: examining the rise of multinational investment in smallholder grain trading in Zambia. *Food Security* **10**, 473-488.

Smale, M., and Tushemereirwe, W. (2007). "An economic assessment of banana genetic improvement and innovation in the Lake Victoria region of Uganda and Tanzania," International Food Policy Research Institute, Washington DC.

Srivastava, A., Dagbenonbakin, G. D., and Gaiser, T. (2010). Effect of fertilization on yam (Dioscorea rotundata) biomass production. *Journal of plant nutrition* **33**, 1056-1065.

Srivastava, A. K., and Gaiser, T. (2010). Simulating biomass accumulation and yield of yam (Dioscorea alata) in the Upper Ouémé Basin (Benin Republic)-I. Compilation of physiological parameters and calibration at the field scale. *Field crops research* **116**, 23-29.

Stone, R. C., and Meinke, H. (2005). Operational seasonal forecasting of crop performance. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **360**, 2109-2124.

Suleiman, A. A., and Ritchie, J. T. (2003). Modeling soil water redistribution during second-stage evaporation. *Soil Science Society of America Journal* **67**, 377-386.

Sun, Y., Solomon, S., Dai, A., and Portmann, R. W. (2006). How often does it rain? *Journal of climate* **19**, 916-934.

Sweeney, S., Ruseva, T., Estes, L., and Evans, T. (2015). Mapping cropland in smallholder-dominated savannas: integrating remote sensing techniques and probabilistic modeling. *Remote Sensing* **7**, 15295-15317.

Syswerda, S. P., Basso, B., Hamilton, S. K., Tausig, J. B., and Robertson, G. P. (2012). Long-term nitrate loss along an agricultural intensity gradient in the Upper Midwest USA. *Agriculture, Ecosystems & Environment* **149**, 10-19.

Tao, F., Rötter, R. P., Palosuo, T., Gregorio Hernández Díaz-Ambrona, C., Mínguez, M. I., Semenov, M. A., Kersebaum, K. C., Nendel, C., Specka, X., and Hoffmann, H. (2018). Contribution of crop model structure, parameters and climate projections to uncertainty in climate change impact assessments. *Global change biology* **24**, 1291-1307.

Teal, R. K., Tubana, B., Girma, K., Freeman, K. W., Arnall, D. B., Walsh, O., and Raun, W. R. (2006). In-season prediction of corn grain yield potential using normalized difference vegetation index. *Agronomy Journal* **98**, 1488-1494.

Tesfaye, K., Gbegbelegbe, S., Cairns, J. E., Shiferaw, B., Prasanna, B. M., Sonder, K., Boote, K., Makumbi, D., and Robertson, R. (2015). Maize systems under climate change in sub-Saharan Africa. *International Journal of Climate Change Strategies and Management*.

Tetteh, F., Larbi, A., Nketia, K. A., Senayah, J. K., Hoeschle-Zeledon, I., and Abdulrahman, N. (2016). "Suitability of soils for cereal cropping in Northern Ghana." IITA.

Tittonell, P., and Giller, K. E. (2013). When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. *Field Crops Research* **143**, 76-90.

Tittonell, P., Vanlauwe, B., de Ridder, N., and Giller, K. E. (2007). Heterogeneity of crop productivity and resource use efficiency within smallholder Kenyan farms: Soil fertility gradients or management intensity gradients? *Agricultural Systems* **94**, 376-390.

Tittonell, P., Vanlauwe, B., Leffelaar, P. A., Shepherd, K. D., and Giller, K. E. (2005). Exploring diversity in soil fertility management of smallholder farms in western Kenya: II. Within-farm variability in resource allocation, nutrient flows and soil fertility status. *Agriculture, Ecosystems & Environment* **110**, 166-184.

Tollenaar, M., Nissanka, S. P., Aguilera, A., Weise, S. F., and Swanton, C. J. (1994). Effect of weed interference and soil nitrogen on four maize hybrids. *Agronomy Journal* **86**, 596-601.

Unganai, L. S., and Kogan, F. N. (1998). Drought Monitoring and Corn Yield Estimation in Southern Africa from AVHRR Data. *Remote Sensing of Environment* **63**, 219-232.

USDA. (2012). "The Yield Forecasting Program of NASS." Retrieved from https://www.nass.usda.gov/Education_and_Outreach/Understanding_Statistics/Yield_Forecasting_Program.pdf

Valdés-Pineda, R., Demaría, E., Valdés, J. B., Wi, S., and Serrat-Capdevilla, A. (2016). Bias correction of daily satellite-based rainfall estimates for hydrologic forecasting in the Upper Zambezi, Africa. *Hydrology and Earth System Sciences Discussions*, 1-28.

Van der Velde, M., Folberth, C., Balkovič, J., Ciais, P., Fritz, S., Janssens, I. A., Obersteiner, M., See, L., Skalský, R., and Xiong, W. (2014). African crop yield reductions due to increasingly unbalanced Nitrogen and Phosphorus consumption. *Global Change Biology* **20**, 1278-1288.

van der Velde, M., and Nisini, L. (2019). Performance of the MARS-crop yield forecasting system for the European Union: Assessing accuracy, in-season, and year-to-year improvements from 1993 to 2015. *Agricultural Systems* **168**, 203-212.

van Ittersum, M. K., van Bussel, L. G. J., Wolf, J., Grassini, P., van Wart, J., Guilpart, N., Claessens, L., de Groot, H., Wiebe, K., Mason-D'Croz, D., Yang, H., Boogaard, H., van Oort, P. A. J., van Loon, M. P., Saito, K., Adimo, O., Adjei-Nsiah, S., Agali, A., Bala, A., Chikowo, R., Kaizzi, K., Kouressy, M., Makoi, J. H. J. R., Ouattara, K., Tesfaye, K., and Cassman, K. G. (2016). Can sub-Saharan Africa feed itself? *Proceedings of the National Academy of Sciences* **113**, 14964.

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., and Lamarque, J.-F. (2011a). The representative concentration pathways: an overview. *Climatic change* **109**, 5-31.

Van Vuuren, D. P., Stehfest, E., den Elzen, M. G. J., Kram, T., van Vliet, J., Deetman, S., Isaac, M., Goldewijk, K. K., Hof, A., and Beltran, A. M. (2011b). RCP2. 6: exploring the possibility to keep global mean temperature increase below 2 C. *Climatic Change* **109**, 95.

Vossen, P. (1990). COMPARATIVE STATISTICAL VALIDATION OF 2 10-DAY WATER-USE MODELS AND OF 3 YIELD-REDUCTION HYPOTHESES FOR YIELD ASSESSMENT IN BOTSWANA. *Agricultural and Forest Meteorology* **51**, 177-195.

Wainwright, C. M., Marsham, J. H., Keane, R. J., Rowell, D. P., Finney, D. L., Black, E., and Allan, R. P. (2019). 'Eastern African Paradox' rainfall decline due to shorter not less intense Long Rains. *npj Climate and Atmospheric Science* **2**, 34.

Wallach, D., Martre, P., Liu, B., Asseng, S., Ewert, F., Thorburn, P. J., van Ittersum, M., Aggarwal, P. K., Ahmed, M., and Basso, B. (2018). Multimodel ensembles improve predictions of crop–environment–management interactions. *Global change biology* **24**, 5072-5083.

Whitbread, A. M., Robertson, M. J., Carberry, P. S., and Dimes, J. P. (2010). How farming systems simulation can aid the development of more sustainable smallholder farming systems in southern Africa. *European Journal of Agronomy* **32**, 51-58.

Willmott, C. J. (1981). ON THE VALIDATION OF MODELS. *Physical geography* **2**, 184-194.

Willmott, C. J., and Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research* **30**, 79-82.

Wright, B., and Cafiero, C. (2011). Grain reserves and food security in the Middle East and North Africa. *Food Security* **3**, 61-76.

Yonah, I. B., Mourice, S. K., Tumbo, S. D., Mbilinyi, B. P., and Dempewolf, J. (2018). Unmanned aerial vehicle-based remote sensing in monitoring smallholder, heterogeneous crop fields in Tanzania. *International Journal of Remote Sensing* **39**, 5453-5471.

You, L., Wood, S., and Wood-Sichra, U. (2009). Generating plausible crop distribution maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. *Agricultural Systems* **99**, 126-140.

Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., Huang, M., Yao, Y., Bassu, S., and Ciais, P. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences* **114**, 9326-9331.

Zheng, J., Mmari, W. N., Nishigaki, T., Kilasara, M. M., and Funakawa, S. (2018). Nitrogen availability to maize as affected by fertilizer application and soil type in the Tanzanian highlands. *Nutrient Cycling in Agroecosystems* **112**, 197-213.

Zheng, J., Qu, Y., Kilasara, M. M., Mmari, W. N., and Funakawa, S. (2019). Nitrate leaching from the critical root zone of maize in two tropical highlands of Tanzania: Effects of fertilizer-nitrogen rate and straw incorporation. *Soil and Tillage Research* **194**, 104295.

Zinyengere, N., Mhizha, T., Mashonjowa, E., Chipindu, B., Geerts, S., and Raes, D. (2011). Using seasonal climate forecasts to improve maize production decision support in Zimbabwe. *Agricultural and Forest Meteorology* **151**, 1792-1799.