The Pennsylvania State University

The Graduate School

DATA-DRIVEN ASSESSMENT OF CLIMATE CHANGE IMPACTS ON AGRICULTURE IN THE SOUTHEAST OF THE UNITED STATES

A Dissertation in
Agricultural and Biological Engineering
by
Hai Nguyen

Submitted in Partial Fulfillment
of the Requirements
for the degree of

Doctor of Philosophy

May 2022

The dissertation of Hai Nguyen was reviewed and approved by the following:

Christine Costello
Assistant Professor of Agricultural and Biological Engineering
Dissertation Adviser
Chair of Committee

Tom L. Richard
Professor of Agricultural and Biological Engineering

Cibin Raj
Assistant Professor of Agricultural and Biological Engineering

Erica A. H. Smithwick
Distinguished Professor of Geography

Allen Thompson
Special Member
Professor Emeritus of Biomedical, Biological & Chemical Engineering
University of Missouri

Suat Irmak
Professor and Department Head of Agricultural and Biological Engineering

Abstract

Agricultural production, in particularly crop production, and food security are already being affected and will continue to face negative consequences of climate change at the global and regional scales. Climate change impacts on crop production are multifold but two aspects remain top priority for researchers: crop productivity and crop water demand. These two aspects need to be examined, taking into consideration the uncertainties in climate change projections and crop modeling approaches. This study is a data-driven regional assessment of climate change impacts on crop yield and crop water demand for major food crops (maize and soybean) in the southeastern United States (U.S.), a less studied but highly populated and lucrative agricultural region.

One of the key challenges in examining the impacts of climate change on the environment and socioeconomic sectors is whether the changes are driven by the background internal climate variability (ICV) or by external, anthropogenic forcings. Therefore, it is important to improve the confidence level of detecting a climate signal from the background ICV when developing adaptation and mitigation strategies, as well as communicating the uncertainties in projecting future climate in certain locations to stakeholders. This study developed a statistical model using a high-resolution mean temperature dataset by Livneh et al. (2015) and a multi-model ensemble of statistically downscaled CMIP5 global climate models (GCM)s for Representative Concentration Pathway (RCP) 8.5 to identify the signal to noise ratio and better understand the climate trend uncertainty due to ICV as it pertains to recent historical climate trends in the conterminous United States (CONUS). A large synthetic observation ensemble was created at a high spatial resolution ($1/16^{th}$ degree) in order to estimate the contribution of ICV to mean temperature variability across CONUS. Then standard bootstrapping procedures was used to estimate variance induced by ICV to estimate the signal-to-noise ratio (SNR). The SNR map revealed which areas that have been or will be experiencing warming trends that is beyond the noise caused by ICV. The results from the large synthetic observation ensemble showed that historically, there is an overall 0.5^{o}C increase in temperature (1961-2010), while under RCP 8.5 scenario, CONUS-mean temperature will increase by a 1.27^{o}C increase in the future (2011-2060). Historical increasing trend in mean temperature was only significant to in the southwestern and some part of southeastern and northeastern regions (SNR > 2). However, the increasing trend will become more prominent in most part in the CONUS for the summer time, where temperature might

increase by 1.52^{o}C (SNR between 4-7). This poses risks of drought and heat exposure for crop production in lower latitude regions. Mean temperature in winter months for the western regions may also experience significant increase (SNR between 2-6) by over 1^{o}C, which may lead to decrease in snow pack in the spring.

Drought is a recurring natural phenomenon that has devastating effects on various economic, social sectors, as well as natural ecosystems and biodiversity. The alarming trend of yield losses to drought in US agriculture has risen over the years, despite increasing financial investments and technological development to abate drought impacts. As various global circulation models project rising temperature and increased precipitation variability in the future, droughts are likely to be exacerbated in terms of frequency and severity, thus increasing the costs and damages in agriculture. The impacts of drought on crop production depend not only on severity, but also the timing of the drought event. Analysis of yield and drought correlation was conducted for two major food crops – maize and soybean – by detrending yield from the USDA-NASS records and modified drought indices SPI and SPEI for at the county level across the southeastern United States (U.S.). The mapping of crop production sensitivity was conducted by combining a quantity of yield loss due to drought and the risk of drought, level of irrigation utility and drought indices. The initial results show that there is a significant negative correlation between yield loss and droughts in most counties. Yield losses (40% on average for maize and 25% for soybean) are significant when abnormally dry conditions occur. Maize is more susceptible to drought than soybean. As about 75% of maize and 64% of soybean are rainfed in the region, the increase in drought under warmer climate will lead to more water needs for crops.

Water supply plays a key role in crop production, as crops require sufficient amount of water to meet the water needs for optimal growth. This study employed a data-driven approach to estimate crop water demand under climate change impacts crop potential evapotranspiration (ET_c) and net irrigation requirement (NIR). Estimations from an ensemble of 18 downscaled GCM that show that there can be a wide range of projected changes in ET_c and NIR as compared to the historical baseline (1980 to 2009). Two scenarios of planting date were considered: the fixed planting date (crops are grown at the historical date) and adaptative planting date (planting date is chosen based on future climate) to examine the impacts of adaptation on NIR. The carbon fertilization effect (CFE) was also considered through two CO_2 level scenarios: the *contemporary*

CO_2 level scenario (378 ppm) and the *elevated* CO_2 level scenario (where future CO2 levels change under RCP 4.5 and 8.5). The study finds that NIR tends to decrease under the fixed planting date scenario compared to the historical baseline due to a shortened growing season, and increase under the adaptive planting date scenario as the climatic demand for evapotranspiration increases due to higher temperature while the risk of accelerated maturity is reduced. The effects of elevated CO_2 level may reduce crop NIR under the same planting date scenario and RCP. Under the same planting date scenario and RCP, if CO_2 level is fixed at the contemporary level, NIR of maize will increase by 5.82 and 3.33 % in the near future and 4.18 and 1.37 % under RCP 4.5 and RCP 8.5, respectively, while NIR of soybean will increase by 13.53 and 4.51 % in the near future, and 9.10 and 2.05 % in the far future under RCP 4.5 and RCP 8.5, respectively. Spatial pattern of NIR shows that the southern area may experience little change or reduction in NIR due to increase in crop water demand under all planting date or NIR scenario as the impacts of increased temperature on shortened growing season length may overcome or cancel out enhanced climatic evaporative demand. The northwestern area near the Corn Belt, where crops were traditionally rainfed, may require investments in irrigation infrastructure, as this area may experience increase in NIR for both maize and soybean.

Understanding climate change impacts on crop production is crucial to anticipating future production and designing appropriate adaptation plans. One of the common approaches for analyzing yield response to climate is to develop empirical models using historical observations of yield and climate. This, however, has a limitation of extrapolating yield response to climate conditions outside the historical range (assumption of stationarity). Emulators, which use simulated climate and yield data from process-based models as 'true' observations to develop statistical models, are used to provide more insight about possible yield responses to climate change. In this paper, two different datasets were used to develop an ensemble of statistical models. The first set of data is historical crop yields and climate data, taken from the USDA-NASS database and gridMET datasets respectively, to construct a set of statistical models based on actual observations (called *observation model*). The second set of data are simulations taken from the Agricultural Model Intercomparison and Improvement Project (AgMIP) to develop an ensemble of *emulators*. These data are comprised of results from the three different global gridded crop models (GGCMs) with climate input from three bias-corrected CMIP5 GCMs with the RCP 8.5. This study focused on estimating of yield changes under future climate scenarios in the

southeastern United States and aiming to analyzes differences in how the emulators simulating yield response to climate compared to the observation model. The emulators and observations show similar direction for the response of yield to changes in temperature and precipitation, but vary in the magnitude. The disparity in yield response to climate as a whole and to each climate component, i.e., precipitation and temperature, between the *observation models* and *emulators* are due to differences in the structure of the GGCMs as well as in the climate data. The results from these models show that under the impacts of climate change, median yield in the study area is expected to decline from 5% to 44% for maize and 3% to 78% for soybean in the near future. For the far future, the decline in yield is expected to be 25% to 72% for maize, and 11% to 96% for soybean.

In conclusion, the study found that climate change poses great risks for crop productivity and water management in the southeastern U.S. There is a risk of yield loss, changes in crop growing cycle and net irrigation requirement. This study utilizes effective and time-efficient data-driven approach to provide useful information for agricultural and water management.

Tables of Contents

Lists of Figures

List of Tables

Acknowledgement

I would not be able to accomplish this degree without the help of many people. First and foremost, I would like to thank my dissertation adviser, Dr. Christine Costello, who has worked relentlessly for me to succeed. Your patience, generosity and intellect have helped me go through this journey, including the most difficult times. I am grateful for your encouragement and support and have enjoyed the past five years of working with you.

I would like to thank my PSU committee members, Dr. Tom Richard, Dr. Cibin Raj, and Dr. Erica Smithwick, for their time and effort spent in reading, editing, guiding, and encouraging along the way. Special thanks to Dr. Allen Thompson, my former co-advisor at the University of Missouri, where I started my doctoral program, who has never given up on me.

Two people who do not serve in my committee, but play an important role in my research, are Dr. Jared Bowden and Dr. Adam Terando of North Carolina State University. Thanks to them, I learned more about the work I am doing and found new areas of interest for my career.

I would like to thank my mother, father and sister for their support, even when this is not what they had envisioned me doing at this stage of my life. Thank you for letting me be who I am and do what I love, even though it hurts you.

Thank you to all of my friends at Penn State and University of Missouri. Thank you for your words of encouragement, coffee breaks, dinners and hangouts. I wouldn't have survived without all of your professional and emotional support, particularly during the pandemic.

The study is part of the NSF-funded project RIPS Type 2 Collaborative Research: Water and Electricity Infrastructure in the Southeast (WEIS), Approaches to Resilient Interdependent Systems under Climate Change - Supplemental Request: Agricultural Sector Interactions with Water and Power Supply Infrastructure under Climate Change - The case of the Southeastern U.S.

Chapter 1 Introduction and Research questions

1.1 *Introduction*

According to the Fifth Assessment Report prepared by the Intergovernmental Panel on Climate Change (IPCC AR5), anthropogenic climate change is occurring and poses great risks for human and natural systems (Field et al., 2014). Greenhouse gas emissions from human activities have continued to rise from the pre-industrial revolution level of ~275 parts per million (ppm) to ~410 ppm as of 2020[1], and is projected to reach 550 ppm by mid-21st century (Earth System Research Laboratories, 2021). Even under the optimistic scenario where nations could achieve net zero emissions targets by 2030, global mean temperature is still set to rise by 1.8^{o}C by 2100 (Climate Action Tracker, 2021), leading to an increase in the frequency and magnitude of extreme weather events (Field et al., 2014). Adverse impacts of climate change on human and natural systems have been observed worldwide, causing disruptions in agricultural production, transportation and infrastructure (Melillo et al., 2014; Vittoz et al., 2013).

Agricultural production and food security are already being affected and will continue to face negative consequences of climate change at the global and regional scales, as evidence around the world show (Zhongming et al., 2019). Change in patterns of precipitation, lengthened dry spells and extreme temperatures are likely to impact crop production through a number of pathways, including shifts in vegetation zones, elevated soil erosion, and intensified photosynthetic rates, etc., (Iizumi et al., 2018; Lobell and Field, 2007; Lobell et al., 2011b; Moore and Lobell, 2015). Numerous studies report an increasing risk of simultaneous crop failure as the global average temperature increases, which would put more pressure on the world economy and human well-being (Gornall et al., 2010; Jin et al., 2017; Lesk et al., 2016; Liebig et al., 2012; Petersen, 2019; Urban et al., 2015; Wheeler and von Braun, 2013). As the population continues to grow and is projected to reach around 10 billion people in 2050 (Department of Economic and Social Affairs, 2019), global food productivity is expected to rise 100–110% compared to the 2005 level to meet the demand (Tilman et al., 2011).

[1] https://www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide

To manage and mitigate the risks of climate change to agriculture, long-term, large-scale and timely strategies play a crucial part. Therefore, agricultural managers, policy-makers and stakeholders must understand the extent of climate change impacts on agricultural productivity to develop effective, efficient management and adaptation plans. With the urgency of climate change threats, there exists a need for robust and timely modeling of climate change impacts on agriculture to assess future risks and corresponding adaptation strategies. Modeling techniques can be generally categorized into: i) process-based, mathematical approach that explicitly simulates various components of agricultural processes and their interactions with climate, environmental and other abiotic factors, often at a daily timestep, and ii) statistical, data-driven approach that concentrates on the functional, reduced-form relationships between agricultural outputs and several variables, including climate (Roberts et al., 2017). Compared to process-based modeling, which requires cumbersome calibration, time and data and human expertise to execute, statistical modeling provides a time-efficient, computationally inexpensive, yet robust and reliable analyses of climate relationships with agricultural outputs (Roberts et al., 2017). For management and planning purpose, statistical modeling would be an efficient and powerful tool, especially in the present-day data abundance (Blanc and Schlenker, 2017).

While the overall consensus is that changes in temperature, precipitation, and carbon dioxide (CO_2) concentration under climate change scenarios are likely to affect agricultural production, the rate, magnitude and pattern of the impacts remain uncertain, particularly at a regional scale, with considerable disagreement across crop models even with consistent climate input data (Ralston Fonseca et al., 2019; Rosenzweig et al., 2017; Rosenzweig et al., 2014; Wilcox and Makowski, 2014). The uncertainty in projecting crop yield response to climate change is the result of differences in climate projection inputs, modelling approaches and spatio-temporal scale (Challinor et al., 2009; Challinor et al., 2014; Hatfield et al., 2011). Given the need to direct decisions in the present for effective and timely adaptation responses, it is critical to develop robust, efficient and rapid methods to assess impacts of climate on agriculture. This research employed an ensemble of statistical models with an ensemble of climate projections from global climate models (GCMs) to analyze climate change impacts on crop production in the southeastern region of the United States (U.S.) to provide reliable and science-based information for the region's management. (describe the SERC). In particular, this research answered some key questions are of great concern amongst scientists and decision-makers, including: *i) can a future*

warming trend during the growing season that may pose significant threats to crop growth be differentiated from the background internal climate variability, ii) how does crop net irrigation requirement change under the future temperature and precipitation pattern, and *iii) what is the direction and magnitude of crop yield change to temperature increase and precipitation pattern changes under future climate scenarios.*

1.2 *Research questions*

This research aims to provide quantitative estimations of climate change impacts on crop yield and crop water demand for major food crops in the SERC. While not focused on the non-climatic factors and their impacts on crop production, this study applies robust data-driven statistical methods, taking into consideration the uncertainties in future climate projections and crop modelling approaches, to provide necessary and insightful information for agricultural and water management of the study region for adaptation and planning. The findings of this study will provide necessary information for strategic planning and further studies on the areas that experience significant changes.

1.2.1 Research question 1

Can a future warming trend during the growing season that may pose significant threats to crop growth be differentiated from the background internal climate variability?

Uncertainties in modeling crop response to climate, first and foremost, came from uncertainties in climate change projections, which are the result of climate model structure, emission scenarios and the internal climate variability (Hawkins and Sutton, 2009). Internal climate variability (ICV) refers to climate variations over time as a result of natural phenomena such as large-scale ocean and atmospheric circulations (e.g.: El Niño-Southern Oscillation). It is an inherent part of the climate system and thus irreducible (Santer et al., 2019). Studies show that ICV can have a large contribution to climate projection uncertainty, and thus makes it difficult for climate change detection and attribution (Barnes et al., 2015; Deser et al., 2010; Deser et al., 2012b; McKinnon and Deser, 2018; McKinnon et al., 2017). Quantifying ICV-induced uncertainty is important to detect whether a climate trend that may pose significant challenges to crop production is emerging, so that decision-makers can avoid any frivolous conclusions on climate change impacts and make strategic adaptation plans. This component of the research lays the foundation for further analyses on climate change impacts on crop production, including crop productivity

and crop water demand. This study employed a recent method by McKinnon and Deser (2018) to develop an large observational ensemble of climate time series in order to quantify ICV-induced uncertainty. An improvement upon previous studies of this type, which mostly used coarse spatial resolutions dataset (> 0.5^{o} gridcell) and at the global scale (Dai et al., 2015; McKinnon and Deser, 2018; McKinnon et al., 2017), is the utilization of high resolution ($1/16^{th}$ degree gridcell) dataset from an ensemble of statistically downscaled GCMs over the conterminous U.S. (CONUS). This study focused on analyzing uncertainty and strength in signal of the warming trend in particular, because: i) uncertainty in future crop yield modeling is dominated by uncertainty in projections of temperature rather than those of precipitation (Lobell and Burke, 2008), and ii) crop yield is generally more sensitive to the changes in temperature, especially when it exceeds the range of what is considered optimal range for crop growth (De Boeck et al., 2012; Jin et al., 2017; Lobell et al., 2011a; Lobell et al., 2013; Schlenker and Roberts, 2009; Stone, 2001; Wahid et al., 2007). There has been very few high-resolution detection and attribution studies, yet these are increasingly important to capture the spatial variability in climate change that is relevant to agricultural management.

1.2.2 Research question 2

How does net irrigation requirement change under the impacts of climate change?

Crop production, as part of the agricultural system, is closely linked to the water infrastructure system, mostly through crop water demand. In the context of integrated water management, it is important to incorporate water demands from all sectors under climate change in a region into planning for adaptation. Most of the food crops in the southeastern U.S. are dominantly rainfed (75% of maize and 62% of soybean), and the area may experience changes in water demands from competitive sectors such as power generation (Cheng et al., 2020b; Cheng et al., 2020a). Therefore, it is crucial to provide robust estimations of crop water demand so that water managers can plan for effective water allocation strategy. This study employed a data-driven approach to estimate crop water demand, taking into account climate change impacts on both the demand side (crop evapotranspiration) and the supply side (precipitation). On the demand side, climate change impacts crop evapotranspiration was estimated as a result of changes in climatic variables (temperature, relative humidity and windspeed, etc.), and also elevated CO_2 concentration (through reduction in stomatal conductance) (Drake et al., 1997; Leakey et al., 2009;

Leakey et al., 2012). Uncertainty in future climate was incorporated through the use of an ensemble of statistically downscaled GCMs. The FAO-56 method of crop evapotranspiration method, including reference evapotranspiration equation and crop characteristic values, were used for rapid and reliable estimations of crop evapotranspiration. Impacts of CO_2 concentration were adjusted by modifying the bulk resistance component in the FAO-56 equation. On the supply side, this study estimated effective precipitation under future climate taking into account soil characteristics (available water storage). Crop net irrigation requirement was estimated as the difference between the supply and demand side. This study provided the ranges of quantitative estimations and spatial distribution of crop net irrigation requirement across the study region for water management and decision-making purpose.

1.2.3 Research question 3

What is the response of crop yields to changes in warming level and precipitation patterns under future climate scenarios?

This study quantified crop yield change as a response to climate change using an ensemble of statistical models. Given the current data availability and statistical techniques, statistical modelling is proven to be more appropriate approach for agricultural management and planning. In recent years, a new method has emerged, which apply statistical methods to outputs from process-based models (Blanc, 2017b; Blanc and Sultan, 2015). However, traditional statistical models are built on historical observations and extrapolate the past relationships between crop and climate to the future (assumption of stationarity), which may exceed the historical records. In recent years, emerged a new approach called *emulators*, in which statistical methods are applied on outputs from a process-based model (Blanc, 2017a; Blanc, 2017b). This new approach has advantages of utilizing outcomes of process-based models, including the simulations of crop response to climate change, thus addressing the assumption of stationarity in traditional statistical models. With increasingly available climate and crop model outputs, such as those implemented under the Agricultural Model Intercomparison and Improvement Project (AgMIP) (Rosenzweig et al., 2013; Ruane et al., 2017), emulators are a more time-effective and less expensive alternative to process-based models. This study utilized AgMIP data to produce an ensemble of emulators, in addition to the traditional statistical model based on historical observations, to assess the possible

range of crop yield responses to climate change for major food crops in the southeastern U.S. This ensemble of statistical models helped to assess the uncertainty in modeling approach.

1.3 *Study region*

The study region of this research is the Southern Electricity Reliability Corporation (SERC), one of the eight regional electric reliability councils under North American Electric Reliability Corporation (NERC) authority. It expands from 91°39N to 75°27N and from 29°37W to 40°40W, comprising 1175 counties from 19 states (Figure 1-1). SERC is home to a large population and lucrative agricultural production (Economic Research Service, 2018), with major food crops including maize and soybean. Investigation on climate change impacts on the food-energy-water nexus examine the potential changes in water and energy requirements is important for the regional management and planning (Ralston Fonseca et al., 2019). However, compared to the Corn Belt and the arid southwestern U.S., this semi-humid region is less studied for climate change impacts on crop production.

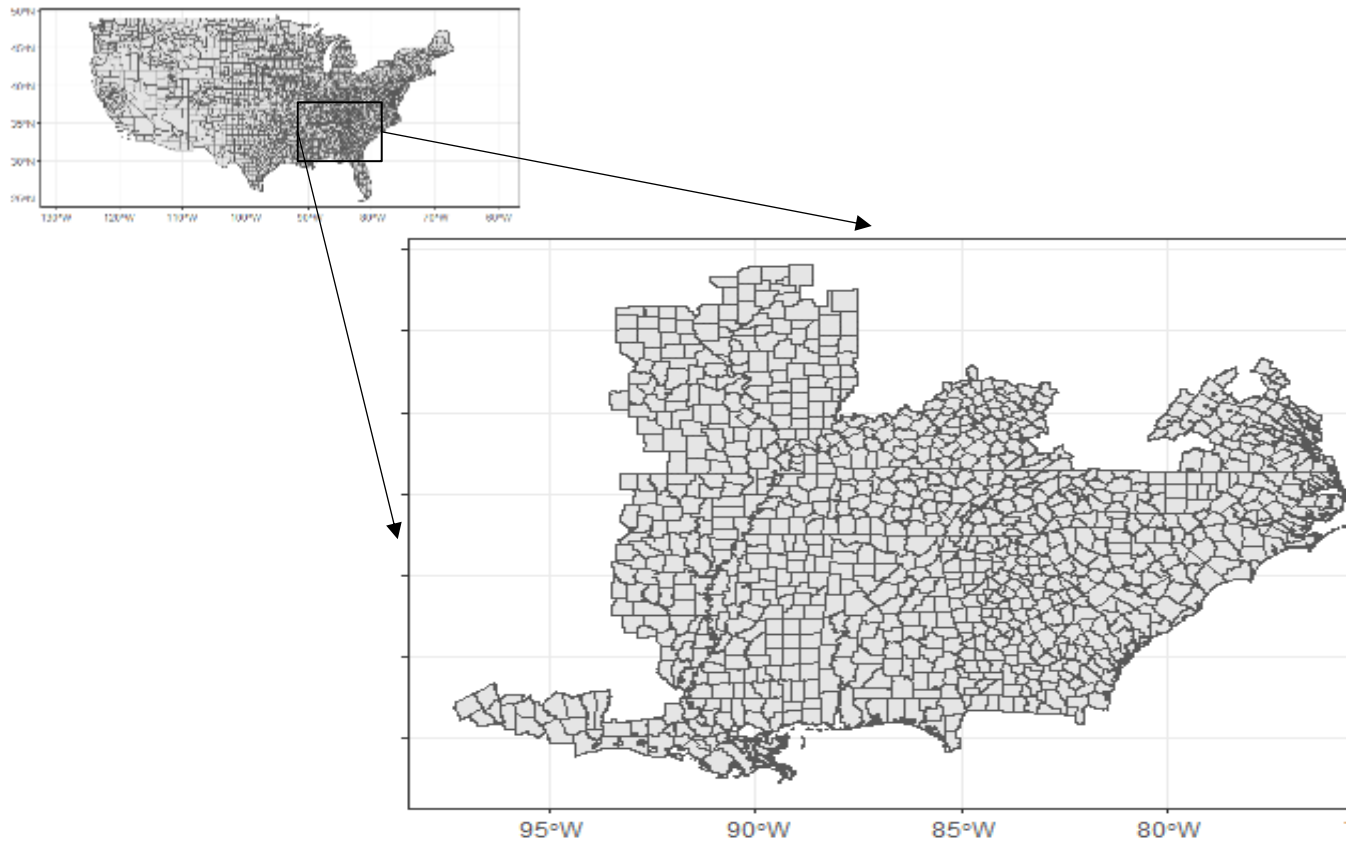

Figure 1-1 Southern Electricity Reliability Corporation (SERC), one of the eight regional electric reliability councils under North American Electric Reliability Corporation (NERC) authority, is located in the southeastern US.

The study region (hereinafter referred to as the SERC) is characterized by the sub-humid climate (see Figure 1-2). Over the growing season of maize and soybean (typically April to September) from 1979-2019, average cumulative precipitation (P) is 648 (±218) mm, while average reference evapotranspiration (ET_o) is 914 (±100) mm. Daily maximum temperature (TX) during the growing season is 30.7 (±2.1) ^oC, while the minimum temperature (TN) is 18.6 (±2.1) ^oC. Daily vapor pressure deficit (VPD) is averaged at 1.06 (±0.26) kPa. Temperature increases from north to south of the SERC, but the southern area is characterized by high precipitation (approximately 900 mm/year), low VPD and ET_o. Most of the climate time series exhibit an upward trend over the historical period, though not significant at $p < 0.05$ (see Appendix A).

1.4 *Document organization*

This dissertation is organized in 6 chapters. Chapter 2 focuses on quantifying the SNR of mean temperature, one of the key climate variables to crop production. The impacts of historical droughts on crop growth in the study region and the vulnerability of the crops to water stress are highlighted in Chapter 3, while Chapter 4 estimates the changes in net water irrigation requirement for the crops in the study region under future climate for regional water management. Chapter 5 answers the question of how much yield may change in the future under climate change scenarios. Finally, chapter 6 concludes the dissertation with a summary of findings and suggestions for further study.

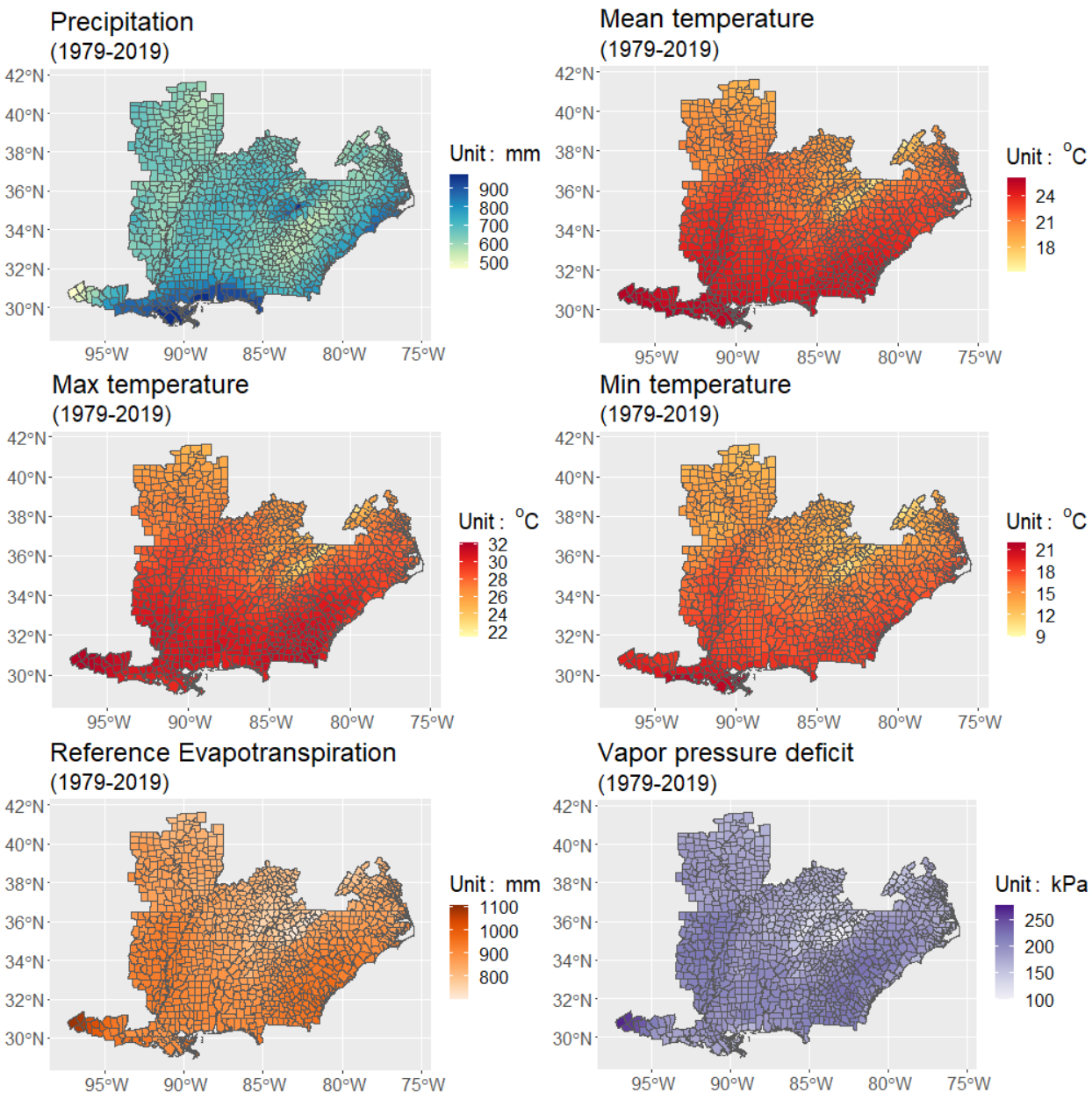

Figure 1-2 Spatial distribution of growing season cumulative precipitation, mean, maximum, minimum temperature, reference evapotranspiration and vapor pressure deficit in the SERC

Chapter 2 Detecting significant trend in mean temperature across the conterminous United States

Abstract

One of the key challenges in examining the impacts of climate change on the environment and socioeconomic sectors is whether the changes are driven by the background internal climate variability (ICV) or by external, anthropogenic forcings. Therefore, it is important to improve the confidence level of detecting a climate signal from the background ICV when developing adaptation and mitigation strategies, as well as communicating the uncertainties in projecting future climate in certain locations to stakeholders. This study developed a statistical model using a high-resolution mean temperature dataset by Livneh et al. (2015) and a multi-model ensemble of statistically downscaled CMIP5 GCMs for RCP 8.5 to identify the signal to noise ratio and better understand the climate trend uncertainty due to ICV as it pertains to recent historical climate trends in the conterminous US (CONUS). A large synthetic observation ensemble was created at a high spatial resolution ($1/16^{th}$ degree) in order to estimate the contribution of ICV to mean temperature variability across CONUS. Then standard bootstrapping procedures was used to estimate variance induced by ICV to estimate the ratio between the signal and the noise (SNR). The SNR map revealed which areas that have been or will be experiencing warming trends that is beyond the noise caused by ICV. The results from the large synthetic observation ensemble showed that historically, there is an overall 0.5^{o}C increase in temperature (1961-2010), while under RCP 8.5 scenario, CONUS-mean temperature will increase by a 1.27^{o}C increase in the future (2011-2060). Historical increasing trend in mean temperature was only significant to in the southwestern and some part of southeastern and northeastern regions (SNR > 2). However, the increasing trend will become more prominent in most part in the CONUS for the summer time, where temperature might increase by 1.52^{o}C (SNR between 4-7). This poses risks of drought and heat exposure for crop production in lower latitude regions. Mean temperature in winter months for the western regions may also experience significant increase (SNR between 2-6) by over 1^{o}C, which may lead to decrease in snow pack in the spring.

2.1 *Introduction*

Climate change poses an imminent threat on the environment and socioeconomic sectors, in particularly agricultural productivity and food security (Hatfield et al., 2011; Rosenzweig et al., 2017). Despite soaring crop yield driven by technological advances and management practices over the last few decades, the agricultural sector remains vulnerable to climate (Lobell et al., 2006; Lobell et al., 2013; Lorenzo-Lacruz et al., 2010; Rosenzweig et al., 2014). Therefore, understanding the impacts of climate change on agriculture is crucial for adaptation planning. Studies on the impacts of climate change in crop production usually aim at quantifying the response of crop yields to the changes in the key climate variables (temperature, precipitation, etc.) either through statistical modelling based on historical observations of crop yields and climate (Lobell and Burke, 2010; Lobell et al., 2006; Schlenker and Lobell, 2010), or by process-based modeling of changes in crop phenology as a result of changes in climate inputs (Acharjee et al., 2017b; Martins et al., 2019b; Rosenzweig et al., 2017; Ruane et al., 2017), or a hybrid method that utilize outputs from process-based models to develop a statistical crop yield model (Blanc, 2017a; Blanc, 2017b; Blanc and Schlenker, 2017; Blanc and Sultan, 2015; Mistry et al., 2017; Schauberger et al., 2017). However, predictions of climate change impacts on crop production are subject to uncertainties from various sources, include crop model structures and climate change projections by global circulation models (GCMs) (Lobell and Burke, 2010; Tebaldi and Lobell, 2008). This part of the research focused on examining uncertainty in climate projections for agricultural modeling.

Uncertainty in climate projections results mainly from to three sources: model uncertainty, and scenario uncertainty and internal climate variability (ICV) (Hawkins and Sutton, 2009). While model and scenario uncertainty can be reduced as understanding of the earth systems and future forcing scenarios improve, ICV will remain inherent to climate change projections. Various studies have been focused at identifying the signal of climate change and estimating the probability of that signal occurring due to natural variability or external forcings (usually referred to as detection and attribution (D&A) studies) (Stott et al., 2010). These studies have revealed the significant influence of ICV on various climate metrics including air temperature, precipitation, etc., undermining the significance in the detected climate trend and making it harder to attributing the observed changes to the driving factors (Barnes et al., 2015, Deser et al., 2010, Deser et al., 2012, McKinnon and Deser, 2018, McKinnon et al., 2017). Therefore, distinguishing a climate trend from the

background ICV is crucial in assessing whether crops are being exposed to shifts in climatic conditions that are truly different from past conditions (Hawkins and Sutton, 2012).

ICV is the natural variability caused by unforced coupled interactions in the climate system between large-scale circulations in the ocean, the atmosphere, the land surface and the cryosphere. Examples include the El Nino Southern Oscillation (ENSO), the Pacific Decadal Oscillation (PDO), the Atlantic Multi-decadal Oscillation (AMO) (McKinnon and Deser, 2018; Dai et al., 2015), of which pattern and magnitude are independent of external forcings such as atmospheric aerosols or greenhouse gas concentrations. ICV can be considered "noise" to the "signal" of the trend in climate variables in the "signal-to-noise ratio" (SNR) when determining the trend of global climate (Santer et al., 2019). If the natural variability is large, the SNR of a climate trend is small, rendering any climate trend detected insignificant. Deser et al. (2016) studied the trend in wintertime surface air temperature (SAT) by combing 30-member initial-condition ensemble of the National Center for Atmospheric Research (NCAR) Community Earth System Model, version 1 (CESM1) and observational datasets over North America. The study found that internal variability contributed 40%–60% of the total warming in the winter time over western Canada and eastern United States and up to 70% in some locations. In addition, removing the ICV almost doubles the SNR of the simulated SAT trends, thus increasing the level of confidence in the detected trends. Since 2000, the rate of increase in the global-mean surface temperature has slowed down despite an increase in the atmospheric greenhouse gas level, which is in sharp contrast with the strong warming rate by model simulations. Dai et al. (2015) analyzed observed and model-simulated global temperature time series since 1920, and found that the Interdecadal Pacific Oscillation (IPO) explains most of the difference between the observed global temperature and the simulated one. The study suggested that IPO could reverse course and lead to accelerated global warming in the coming decades. Deser and Wettstein (2014) found that the trends in September Arctic sea ice extent during 2020–2059 range from -2.0×10^6 to -5.7×10^6 km^2 across a 40-member Community Climate System Model, version 3 (CCSM3) ensemble for the period 2000–2061, indicating a substantial role of internal variability in future Arctic Sea ice loss projections. Similarly, a three-fold range (from -7.0×10^3 to -19×10^3 km^3) is found for summer Artic sea ice volume trends.

Many D&A studies rely on large ensembles of global circulation model (GCM) experiments over a sufficiently long time period (Stott et al., 2010; Laepple and Huybers, 2014; Deser and Wettstein, 2014; Hawkins and Sutton, 2012). Examples are shown in Figure 1 (Stott et al., 2010), where a 'control' simulation of a climate variable over a period of time with no changes in the external drivers of climate does not show any the sustained trend in temperatures as observed in the historical records (right panel). In contrast, the simulations from an ensemble of GCMs under changes in anthropogenic forcings (such as increases in the atmospheric greenhouse gas (GHG) level) tend to be consistent with trends in the observed records (left panel). Comparison of observed changes in climate variables to the simulations of GCMs can be used to produce the likelihood of which the changes could occur due to natural variability. In addition, using a large ensemble of GCMs also is an increasingly common method in evaluating the contribution of anthropogenic forcings and ICV to the trends of climate variables (Laepple and Huybers, 2014; Deser and Wettstein, 2014; Barnes et al., 2015; Dai et al., 2015; McKinnon and Deser, 2018; McKinnon et al., 2017). Two common types of large ensembles used in D&A studies are initial-condition ensembles and multi-model ensembles. An initial-condition ensemble consists of experiments from a single GCM under the same forcing scenario, but slightly different initial conditions for each member (Kay et al., 2015). A multi-model ensemble includes multiple models which is run under the same forcing scenario, such as ensemble from phase 5 of the Coupled Model Intercomparison Project (CMIP5) (Taylor et al., 2012). The spread in the simulations of the ensemble, which is the result of internally generated variability by the individual ensemble member, is considered to present the range of possible variability caused by simulated internal variability (Stott et al., 2010; McKinnon et al., 2017). This approach, however, is limited by the ability of the model ensemble to accurately simulate the spatiotemporal covariance of the real-world climate, which is generally not true. To counter this issue, McKinnon et al. (2017) applied a statistical resampling methods to gridded observational winter near-surface temperature records over North America for a 50-year period (1966-2015) in order to quantify uncertainty due to ICV, and concluded that in comparison to uncertainty in simulated trend from National Center for Atmospheric Research (NCAR) Community Earth System Model version 1 (CESM1) Large Ensemble tend to overestimate the ICV. Barnes et al. (2015) argued that argued that the uncertainty in future climate trends due to internal variability can be robustly estimated from the statistics of the observed climate.

(1) **(2)**

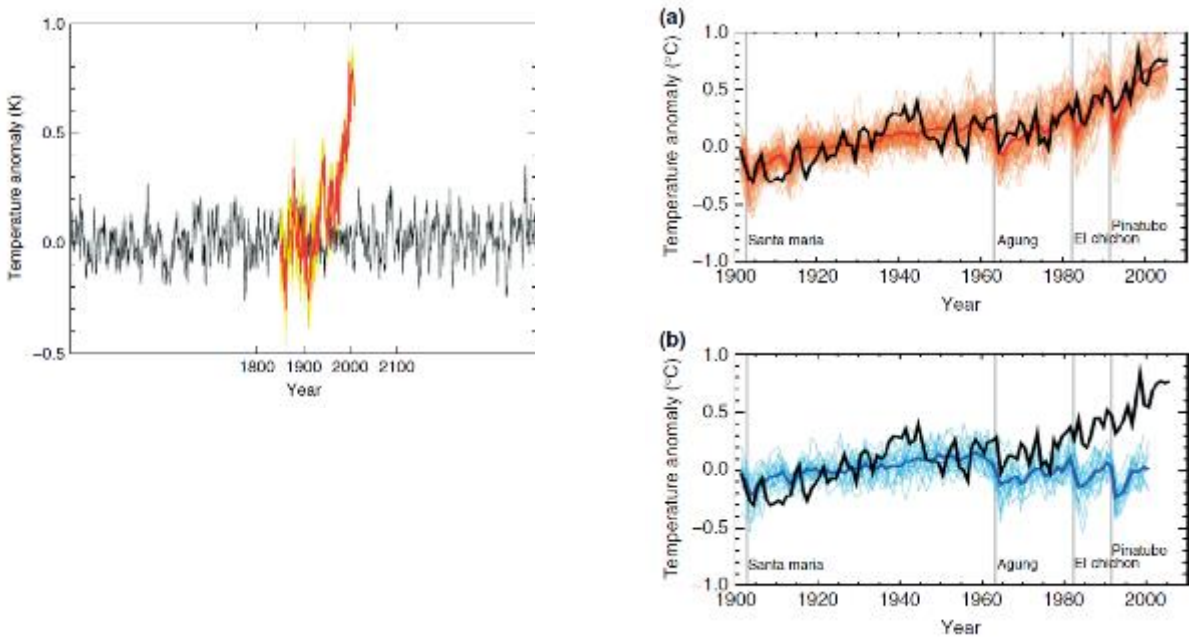

Figure 2-1 Examples of detection and attribution studies. (Left panel) Observed global mean temperature anomalies from 1850 to 2008 (in red) relative to the mean temperature over the 1861–1899 period superimposed on a 1000-year "control run" simulation of global mean temperatures from the HadGEM1 model (black line). (Right panel) Observed global mean surface temperature anomalies during 20th century relative to the period 1901–1950 (black line) compared to simulations from an ensemble of GCMs from the IPCC AR4 report with (a) anthropogenic forcings (red lines) and (b) natural forcings only (blue lines). Vertical gray lines indicate the timings of major volcanic eruptions. The thick lines show the ensemble means and the thin lines show individual simulation results (extracted from Stott et al. (2010)

A series of published studies have introduced new methods and accompanying datasets to conduct climate signal detection based on a large ensemble of model runs from a single GCM (McKinnon et al., 2017; McKinnon and Deser, 2018; Lehner et al., 2017; Barnes et al., 2015). These methods rely on: i) time series analysis of trend uncertainty (Barnes et al., 2015), and ii) use of large ensemble GCM (either multi-model (Dai et al., 2015) or initial-condition ensembles (McKinnon et al., 2017; McKinnon and Deser, 2018) to separately estimate ICV and anthropogenic forcings contributions to climate trends over the past 50-100 years. McKinnon et al. (2017) and McKinnon and Deser (2018) proposed an alternative approach in which a statistical model is developed to generate multiple sequences of climate simulations within the constraint of the statistics (mean and variance) of the observational records. The variability in the climate time series generated by the statistical model is analogous to the climate fluctuations generated by perturbing the initial condition in an initial-condition ensemble.

However, most D&A studies have been conducted at a very coarse spatial scale (> 500km); very few studies focus on analyzing the changes at finer spatial resolution at the regional level, even though changes can impact agricultural crops and livestock at a much smaller scale (4 - 20km gridcell). Regional D&A studies face some of the key challenges. First, large natural variability in regional climate, especially at higher latitudes, lead to high level of ICV-induced noise that can overwhelm the signal of anthropogenic forcing in some locations, thus reducing the confidence level of detected climate trend in some regions (Shindell and Faluvegi, 2009; Hawkins and Sutton, 2012). Secondly, there is uncertainty in GCMs to reliably capture the effects of anthropogenic and natural forcings on climate at the regional level (Stott et al., 2010). And finally, the dearth of long-term and fine resolution climate data that for sampling of long-term large-scale circulations such as ENSO or PDO.

In this study, the SNR of mean temperature across the conterminous United States (CONUS), in particular the southeastern U.S., is estimated using the method proposed in McKinnon and Deser (2018) to generate a 1000-member synthetic observational large ensemble (OBSLE) of temperature. A statistical model that incorporate the anthropogenic forcings and the modes of large-scale earth circulations was developed. This study focused mainly on mean temperature for several reasons: i) uncertainty in future crop yield modeling is dominated by uncertainty in projections of temperature rather than those of precipitation (Lobell and Burke, 2008), ii) crop yield is generally more sensitive to the changes in temperature, especially when it exceeds the range of what is considered optimal range for crop growth; and iii) current D&A studies that explore changes in precipitation in the CONUS do not find consistently significant signals (Guo et al., 2019; Knutson and Zeng, 2018; McKinnon and Deser, 2018). Studies for precipitation requires more detailed data that represent more closely the physical processes of precipitation via dynamical downscaling, which is not available at the moment for this spatial resolution scale.

The goal of this study was to quantify the contribution of ICV to mean temperature trends over the period of 1961-2010 in CONUS and the strength of the temperature signal in the future period (2011-2060) at the monthly timestep, with an emphasis on the growing season of maize and soybean (generally from March to October across the CONUS, and April to September for SERC in particular). In addition, this study highlighted the differences in SNRs as a result of differences

in data resolution by generating two 1000-member ensembles using two different datasets: i) coarse global gridded observational temperature data coupled with 40-member ensemble of the large-scale model NCAR CESM1 (CESM1-LE) at a grid resolution of 0.5°x0.5° , and ii) fine 1/16th degree gridded temperature data with 20-member multi-model GCMs that are statistically downscaled to the CONUS by Abatzoglou and Brown (2012). This study complements a D&A study on climate change impacts on biologically-relevant climate indices by the National Climate Adaptation Science Center (NCASC).

2.2 *Methodology and data*

1.1 Statistical model

McKinnon and Deser (2018) proposed a method that generate alternate sequences of climate variability that is consistent with the spatio-temporal mean, variance and autocovariance of the observation records (Figure 2-2). At each grid box, a linear regression is fit for the observational temperature as a combination of its mean state, its response to external forcings, and its response to the three large-scale modes:

$$X^{i,t} = \beta_0^i + \beta_F^i * F^t + \sum_{m=1}^{3} \beta_M^{i,m} * M^{m.t} + \varepsilon^{i,t} \qquad \textit{Eq. 2-1}$$

Where *i, t* and *m* denote grid, time and one of the three large-scale modes, respectively. *X*: climate variable (mean temperature), *F:* forced trend, *M:* modes of SST variability by one of the three large-scale circulations (ENSO, PDO, and AMO), ε: residual. Coefficients β_0 is the mean state of the climate variable, while β_F , β_M represents the sensitivity of the climate variable to the external forcing and the large-scale SST modes, respectively. The coefficients were estimated by ordinary least squares.

Time series of the external forcing influence F^t was calculated for each of the 12 months using the Dai et al. (2015) method. The method assumes that the global-mean, ensemble-mean time series of a given variable in a climate model ensemble can be used to represent the temporal structure of the climate response to external forcings (Tebaldi and Knutti, 2007). In this study, the forced trend was estimated at both the global scale and at the regional scale to examine for the uncertainty in the models' simulation of regional response to external radiative forcings. Global F^t was estimated from the CESM-LE ensemble during the 1920-2014 period, and CONUS F^t was estimated from both CESM-LE and MACA-L13, with the latter being constrained to the 1950-

2011 period due to data availability. In order to account for the model bias in simulations of the warming rate (Dai et al., 2015), all the time series F^t is detrended from their long-term mean and represent the year-to-year variations in the forced trend using the Climate Variability and Diagnostic Package (Phillips et al., 2014). Since high level of warming presents a higher risk to agriculture, the radiative concentration pathway (RCP) used in this study is RCP 8.5.

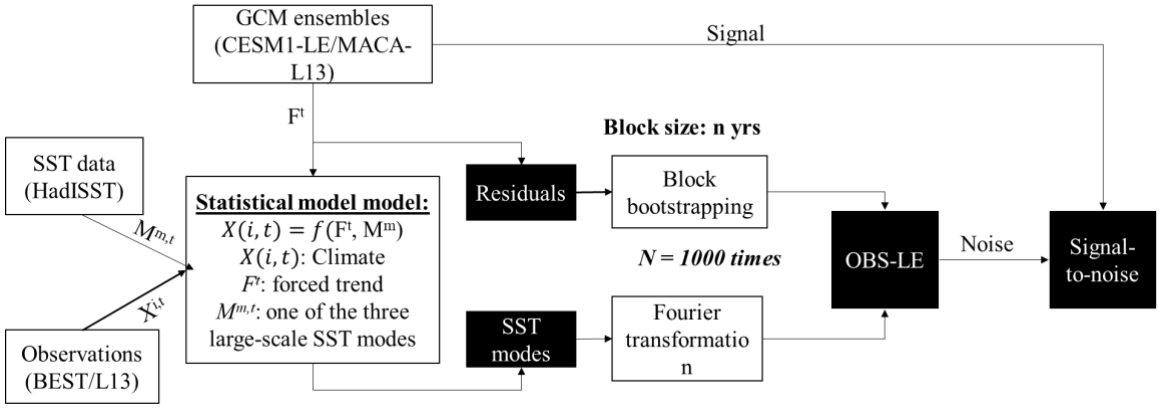

Figure 2-2 Diagram of methodology to quantify the internal climate variability (ICV) according to (McKinnon and Deser, 2018; McKinnon et al., 2017). The indices i, t and m denote the grid box, year and mode of SST. The three modes of SST were estimated by regressing SST data from the HadISST dataset (Rayner, 2003) against the observational climate dataset of three large-scale ciculations PDO, ENSO and AMO. The signal is estimated as the magnitude of trend in the ensemble-mean, region-mean temperature

The component $M^{m,t}$ is one of three orthogonalized SST time series estimated from principal component analysis of the original observed ENSO, PDO and AMO time series to eliminate the potential collinearity between the three modes. The SST pattern associated with each mode is estimated by regressing SST data from the HadISST dataset (Rayner, 2003) against the observational climate dataset, using the Climate Variability and Diagnostic Package.

2.2.1 Observational data and climate model simulations

Monthly temperature observations for the CONUS are developed by Livneh et al. (2013) data provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, from their website[2]. The Livneh dataset (herein after referred to as L13) is gridded at a spatial resolution of 1/16th degree resolution, derived from daily temperature observations from approximately 20,000 Cooperative Observer (COOP) stations operated by the National Ocean and Atmospheric Agency (NOAA),

[2] https://psl.noaa.gov/data/gridded/data.livneh.html

using the synergraphic mapping system (SYMAP) algorithm (Livneh et al., 2013). The mean temperature time series over CONUS is averaged from the monthly maximum and minimum temperature datasets. In addition, mean temperature data are also taken from the Berkeley Earth Surface Temperature (BEST) 1^{o} gridded dataset, created through spatial and temporal interpolation (kriging) of in situ station measurements that covers the earth land surface (Rohde et al., 2013).

For the three large-scale earth circulations, monthly indices of the modes of sea surface temperature (SST) variability were used: Niño-3.4 averaged over 10^{o}N–10^{o}S, 170^{o}–120^{o}W, for El Niño–Southern Oscillation (ENSO)[3]; the leading principle components of monthly SST anomalies in the North Pacific Ocean, poleward of 20N as standardized values for the Pacific decadal oscillation (PDO) [4]; and the average of SSTs over the North Atlantic (0^{o}–80^{o}N) to represent the Atlantic Multidecadal Oscillation (AMO)[5]. The datasets for the PDO and AMO indices have been subtracted from their respective global means.

To represent the response of the climate system to the anthropogenic forcings, two ensembles of models were used in this study. For the CONUS, a multi-model ensemble of 20 statistically downscaled CMIP5 models to 1/16th degree gridcell using the Multivariate Adaptive Constructed Analogs (MACA) method by Abatzoglou (2013) with L13 dataset for the period 1950-2013 as the bias-correction dataset (hereinafter referred to as the MACA-L13) was used. Note that the MACA-L13 ensemble is constrained by the data availability, which only starts from 1950 to the end of the century. The forced response is estimated as the CONUS-wide, ensemble-wide mean across all 20 models in the ensemble. An initial-condition ensemble of 40 members from the Community Earth System Model (CESM) model (CESM-LE) was also used to estimate the global and CONUS-specific forced trend. Each member of the ensemble was simulated by perturbing the initial atmospheric conditions on 1 January 1920 of a single parent simulation that begins in 1850 (Kay et al., 2015). These CESM-LE estimates served as comparisons to highlight the regional variability in detecting the signal to noise ratio.

[3] https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/nino34.long.data
[4] http://research.jisao.washington.edu/pdo/PDO.latest.txt
[5] https://www.esrl.noaa.gov/psd/data/correlation/amon.us.long.data

2.2.2 Generating synthetic observational ensembles (OBS-LE)

The model in Eq. 2-1 is capable of removing most of the forced trend from the observational records, and thus revealing the unforced pattern of the ICV, that is the large-scale modes and the climate noise (residuals) from the model. To estimate the contribution of ICV to the uncertainty in the strength of climate signal over the years, the SST modes and the residual from the model can be used to produce alternate time series of the temperature observations that could have occurred due the stochastic nature of the coupled ocean-atmospheric circulations. This method is analogous to the perturbation and propagation of initial conditions in the CESM1-LE.

McKinnon and Deser (2018) made a couple of assumptions for the surrogate time series. First, both the SST modes and residual are stationary, meaning their variability remain unchanged over time. Second, SST modes do not have initial-value predictability which may reduce the spread of each mode at the beginning of the modeled time period. Both of these assumptions are based on the current observations that there have been no changes in the modes, and there is no strong evidence that these modes may respond to the aerosol changes over the next 50 years (Chen et al., 2017; Murphy et al., 2017; Okumura et al., 2012; Smith et al., 2016). As the SST modes and the residual have different temporal structure, different approaches were applied to create alternate time series for each of them. The SST mode time series were produced by transforming the original data into the Fourier phases, and multiplying them by uniformly distributed random phases before transforming back into a time series. By multiplying the new time series with the regression coefficients $\beta_M^{i,m}$, the spatiotemporal patterns of the temperature records were reproduced, equivalent to sampling of the ENSO, PDO, and AMO time series. For the residual time series, a block bootstrap procedure was applied by grouping the time series into time blocks of 2 years, randomly resampling these blocks with replacement. This approach allowed the full spatial structure of the anomalies in the climate variable to be replicated in the surrogate data. Both the randomization of the mode time series and the block bootstrap are performed 1000 times to produce a synthetic observational ensemble of climate variable data (referred to as OBS-LE).

The variability in the trend of each member of the OBS-LE represented the variability in the climate predictions caused by ICV. However, it may not represent the "true" response of the climate variable to anthropogenic forcings like the climate model ensembles. Therefore, for estimates of the trend of temperature during the historical period (1961-2010) and future period

(2011-2060), estimates from the ensembles (CESM1-LE and MACA-L13) were used instead. The synthetic observational ensembles created by the model in Eq (1), therefore, were recentered to the forced trends of the corresponding model ensemble. In this study, two OBS-LE were created, and referred to as OBS-MACA (using L13 and MACA-L13) and OBS-CESM1 (using BEST and CESM1-LE).

2.2.3 Quantifying the strength of the signal in climate predictions

In order to quantify the contribution of ICV to uncertainty over the spatial extent of CONUS, the signal-to-noise ratio (SNR) was estimated at each grid box. The SNR is defined as the ratio between the magnitude of climate trend over the period of 50 years (signal) and standard deviation in all the possible trends (noise). The trend of temperature over the historical period (2061-2010) and the future period (2011-2060) were estimated as the magnitude of change in the model ensemble mean (MACA-L13 and CESM1) under RCP 8.5, while the noise is estimated as the standard deviation of all the trends estimated from the 1000 members of their corresponding surrogate OBS-LE (OBS-MACA and OBS-CESM1). This allowed the whole range of possible variability in the climate sequence caused by the different sampling of the internal variability in the Earth's coupled ocean-atmosphere circulations to be taken into consideration, giving a robust estimation of uncertainty. Since it was assumed that the variability in the SST and residual of the model in Eq. 2-1 remain stationary over the entire simulation period of climate models, the noise was the same for both the historical and future periods. Therefore, the SNR reflected whether a signal in climate trend has become larger over the years due to external climate forcings. The higher the SNR, the stronger the signal that a trend has emerged from the "background" noise caused by the natural variability. In general, SNR exceeding 2 indicates a significant trend.

2.3 *Results and discussion*

2.3.1 Regional variability of forced trend due to anthropogenic forcings

There has been discussion about the uncertainty in GCMs' simulations of regional response to radiative forcings such as greenhouse gas emission (Adachi et al., 2018; Deser et al., 2012a; Hawkins and Sutton, 2010). By superimposing the forced trend estimated from the MACA-L13 for CONUS (referred to as CONUS-L13), and the CESM1 for CONUS and the global mean temperature (referred to as CONUS-CESM1 and GLOBE-CESM1, respectively), the increase in variability of the response of temperature to climate forcings was revealed. While all the three

forced trends showed similar direction and magnitude of change, they show different level of temporal and ensemble variability (Figure 2-3 and Table 2-1). The GLOBE-CESM1 trend showed small variability across all months both in terms of temporal variability and across the ensemble. However, the CONUS-CESM1 showed increased variability in time and across the ensemble compared to the global mean estimations from the same ensemble of GCMs. The variability in the winter months in the CONUS-CESM1 trend increased by two to four times compared to that of the global trend, while the summer months had a smaller spread, but still doubled the spread in the global trend (Figure 2-4). The increase in regional variability of GCM predictions have been discussed in previous studies (Stott et al., 2010; Hawkins and Sutton, 2010; Deser et al., 2012a). The intensified variability is attributable to: i) that regional temperature and precipitation are affected by low-frequency variability in large-scale coupled ocean-atmospheric circulations, such as ENSO, PDO and AMO, while this effect is canceled out on a global scale (Stott et al., 2010), ii) the importance of factors, such as land use change, irrigation or reservoir construction, etc. which could magnify at a smaller scale than the global scale (Pielke et al., 2011; Lobell and Bonfils, 2008), and iii) limitations of GCMs in capturing characteristics of regional climate variability (Ban et al., 2015; Adachi et al., 2018). When using statistically downscaled multi-model GCMs such as the MACA-L13 ensemble, the variability in climate trend predictions increases further due to the differences in model structure and increase in spatial variability.

Table 2-1 Estimated trends (°C/year) and standard deviation (SD over the period 1950-2011 (statistics for the 1920-2017 in the brackets) estimated from the CESM1 and MACA-L13 ensembles for CONUS and the global mean temperature. All the trends are significant at $p < 0.001$

Months	**CONUS_MACAL13**		**CONUS_CESM1**		**GLOBE_CESM1**	
	Trend	**SD**	**Trend**	**SD**	**Trend**	**SD**
January	0.0112	1.87	0.0137 (0.0094)	1.39 (1.42)	0.0113 (0.0079)	0.30 (0.33)
February	0.0149	1.90	0.0139 (0.0089)	1.33 (1.33)	0.0116 (0.0080)	0.30 (0.33)
March	0.0145	1.61	0.0141 (0.0089)	1.03 (1.03)	0.0117 (0.0081)	0.29 (0.32)
April	0.0142	1.26	0.0128 (0.0081)	0.68 (0.70)	0.0117 (0.0081)	0.27 (0.31)
May	0.0143	0.97	0.0136 (0.0081)	0.49 (0.51)	0.0116 (0.0079)	0.27 (0.30)
June	0.0193	0.88	0.0142 (0.0081)	0.49 (0.50)	0.0116 (0.0078)	0.27 (0.30)
July	0.0208	0.80	0.0133 (0.0077)	0.47 (0.48)	0.0111 (0.0076)	0.26 (0.29)
August	0.0220	0.82	0.0125 (0.0076)	0.46 (0.48)	0.0106 (0.0074)	0.24 (0.28)
September	0.0220	0.93	0.0139 (0.0086)	0.50 (0.52)	0.0114 (0.0079)	0.26 (0.29)
October	0.0199	1.11	0.0115 (0.0075)	0.59 (0.62)	0.0127 (0.0087)	0.29 (0.32)
November	0.0193	1.45	0.0123 (0.0078)	0.89 (0.90)	0.0128 (0.0087)	0.30 (0.34)
December	0.0173	1.77	0.0152 (0.0087)	1.24 (1.26)	0.0122 (0.0083)	0.30 (0.33)

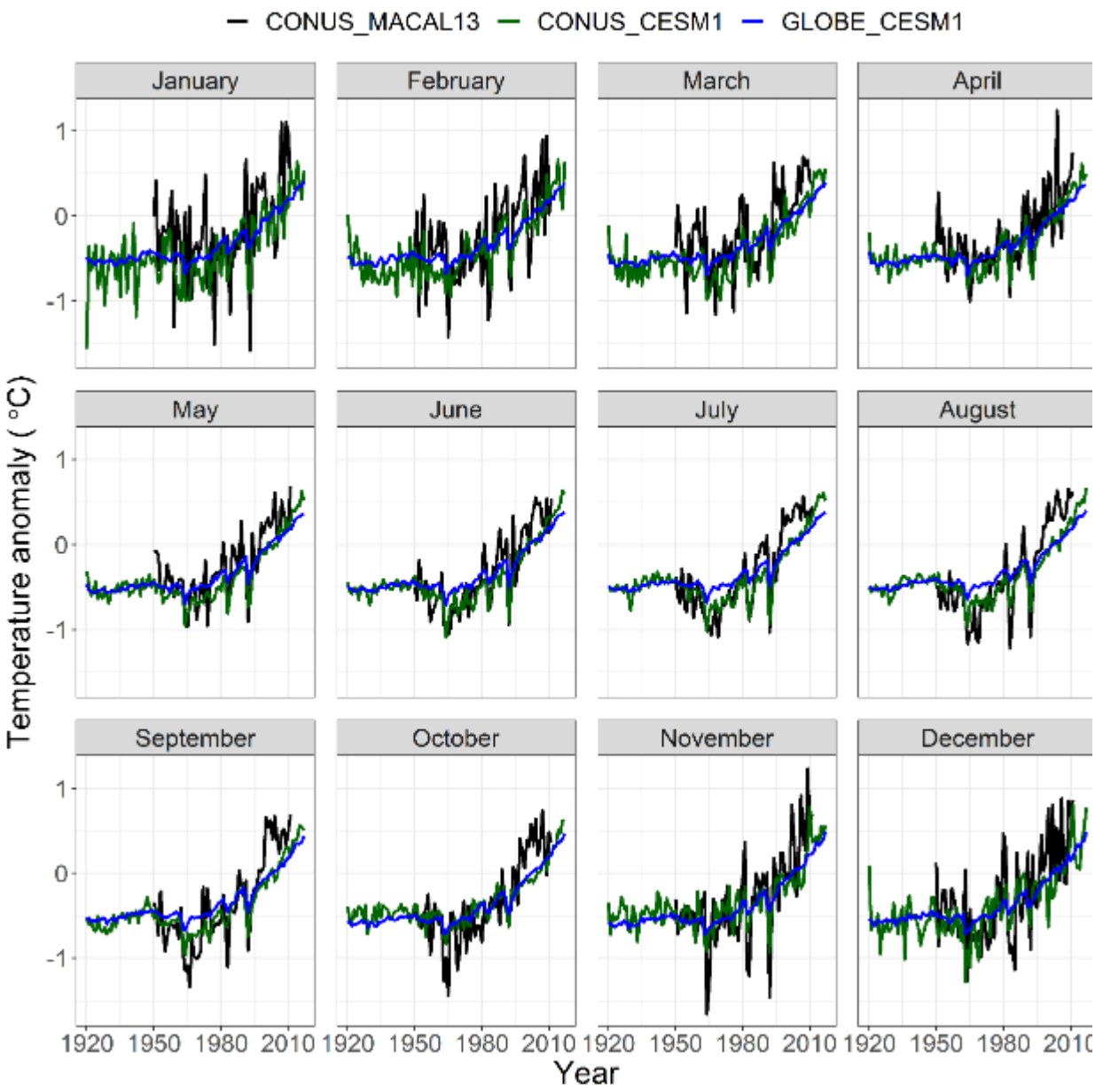

Figure 2-3 The forced trend of mean temperature anomalies (°C) for each month derived as the ensemble mean, region-mean from the CESM1-LE and the MACA-L13 ensemble. Temperature simulations from CESM1-LE were used to estimate the global mean (blue line) and CONUS (green line), while simulations from MACA-L13 were used to estimate the CONUS-wide trend (black line)

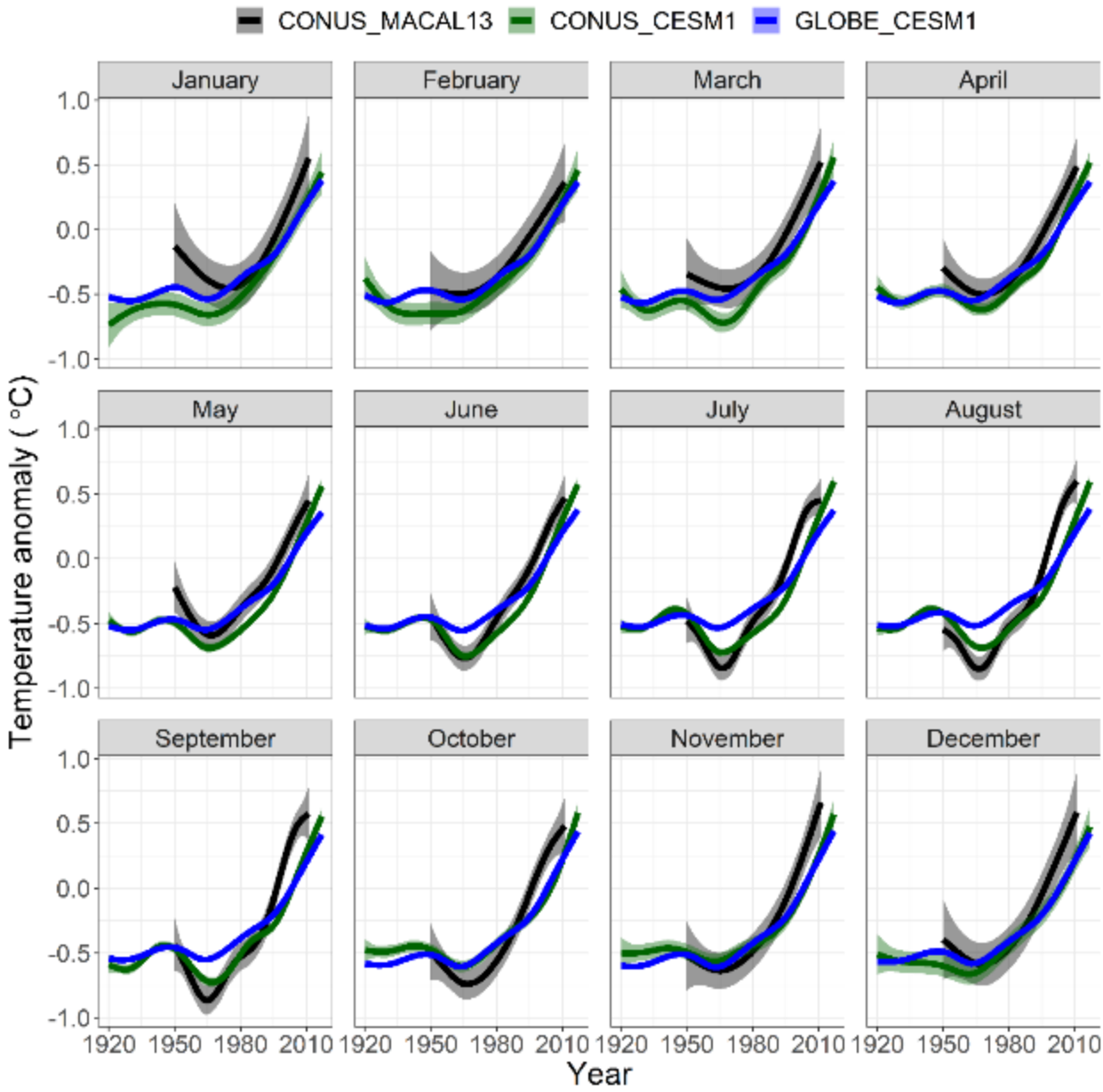

Figure 2-4 The forced trend of mean temperature anomalies (°C) for each month, smoothed from the simulations from the CESM1-LE and the MACA-L13 ensemble using generalized additive model (GAM). Temperature simulations from CESM1-LE were used to estimate the global mean (blue line) and CONUS (green line), while MACA-L13 was used to estimate the CONUS-wide trend (black line). The responding colored shading s represent the 95% confidence level of the forced trend as estimated from the ensembles

In general, the forced trend shows a significant upward trend period ($p < 0.001$) in mean temperature for all months during the 1950-2011 (see Figures 2-4 and Table 2-1) across all the climate projections. The rates of temperature increased in CONUS during the winter months (October to March) were similar between CONUS-L13 and CONUS-CESM1. During the crop growing season (April to September), however, CONUS-MACAL13 showed a significantly

higher rate of increase, almost twice as much in August, compared to the estimates from CONUS-CESM1 and GLOBE-CESM1, especially since 1990s to present day. However, the variability in the trend was also larger for the CONUS-MACAL13, followed by the CONUS-CESM1, while the global estimates have the smallest variability. The variability increased greatly for the winter months, compared to the growing season, for CONUS trend estimates from both of the ensembles. The differences in the rate of change and variability of the trend between the three ensembles showcased the increased uncertainty in response to climate forcings at the regional level, highlighting the needs for better understanding of modeling the climate system especially at the smaller scale.

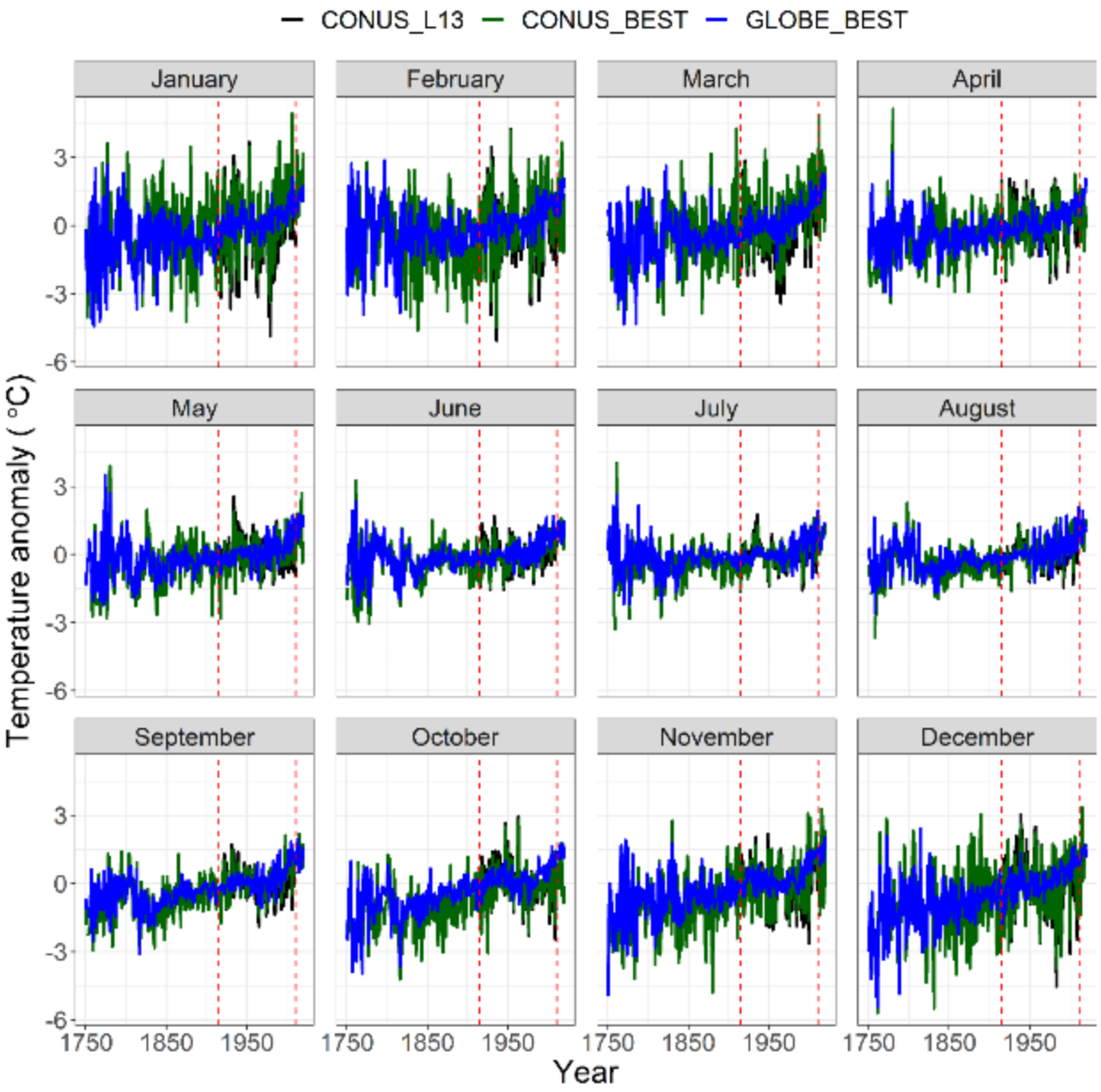

Figure 2-5 The observational mean temperature anomalies (°C) for each month derived as region-wide average from the L13 and the BEST dataset. The global mean (blue line) and CONUS mean (green line) were estimated from BEST dataset from 1750 to 2020, while L13 data were used to estimate the CONUS-wide (black line) time series from 1950-2011 (framed by the two red dashed line)

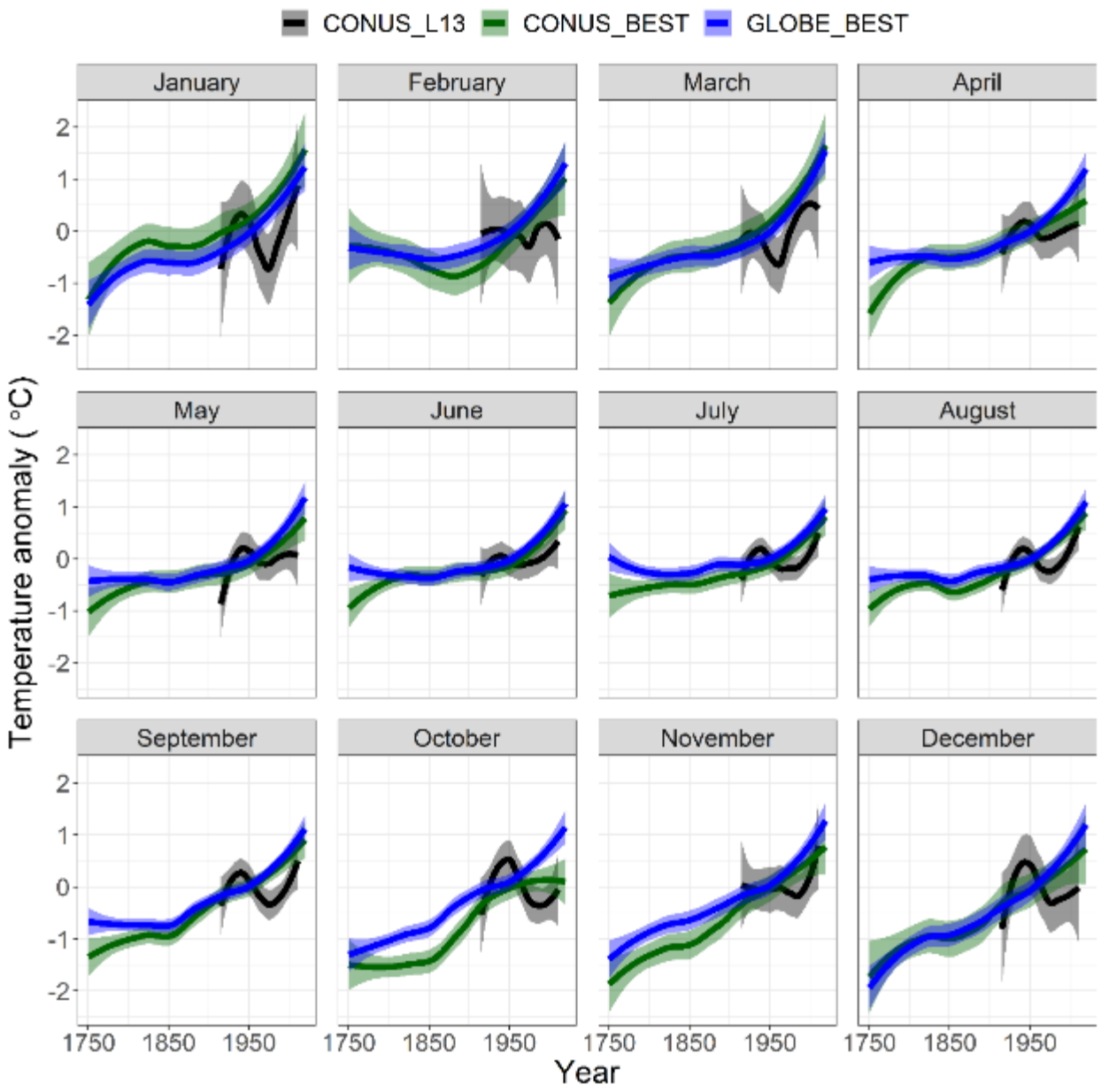

Figure 2-6 The observational mean temperature anomalies (°C) for each month smoothed from the L13 and the BEST datasets for the CONUS (black and green lines) and the global mean (blue line). The shaded areas represent the standard deviation of the mean temperature trend.

While evidence of the global warming trend is clear as shown in this study as well as previous studies, regional warming trends are usually overwhelmed by the ICV (Barnes et al., 2015; Hawkins and Sutton, 2009; McKinnon et al., 2017). The mean temperature trends of CONUS and the world using the observational records (CONUS-L13 estimates from the L13 dataset, and the CONUS-BEST and GLOBE-BEST estimates from the BEST dataset) incorporates both the forced trend and the internal variability (Figures 2-5 and 2-6). CONUS-L13 has higher level of variability and lower increase in trend, especially in the winter months. This shows that

the influence of ICV could mask the impact of climate change in the region, especially in at a smaller scale.

2.3.2 Magnitude and variability in temperature trends

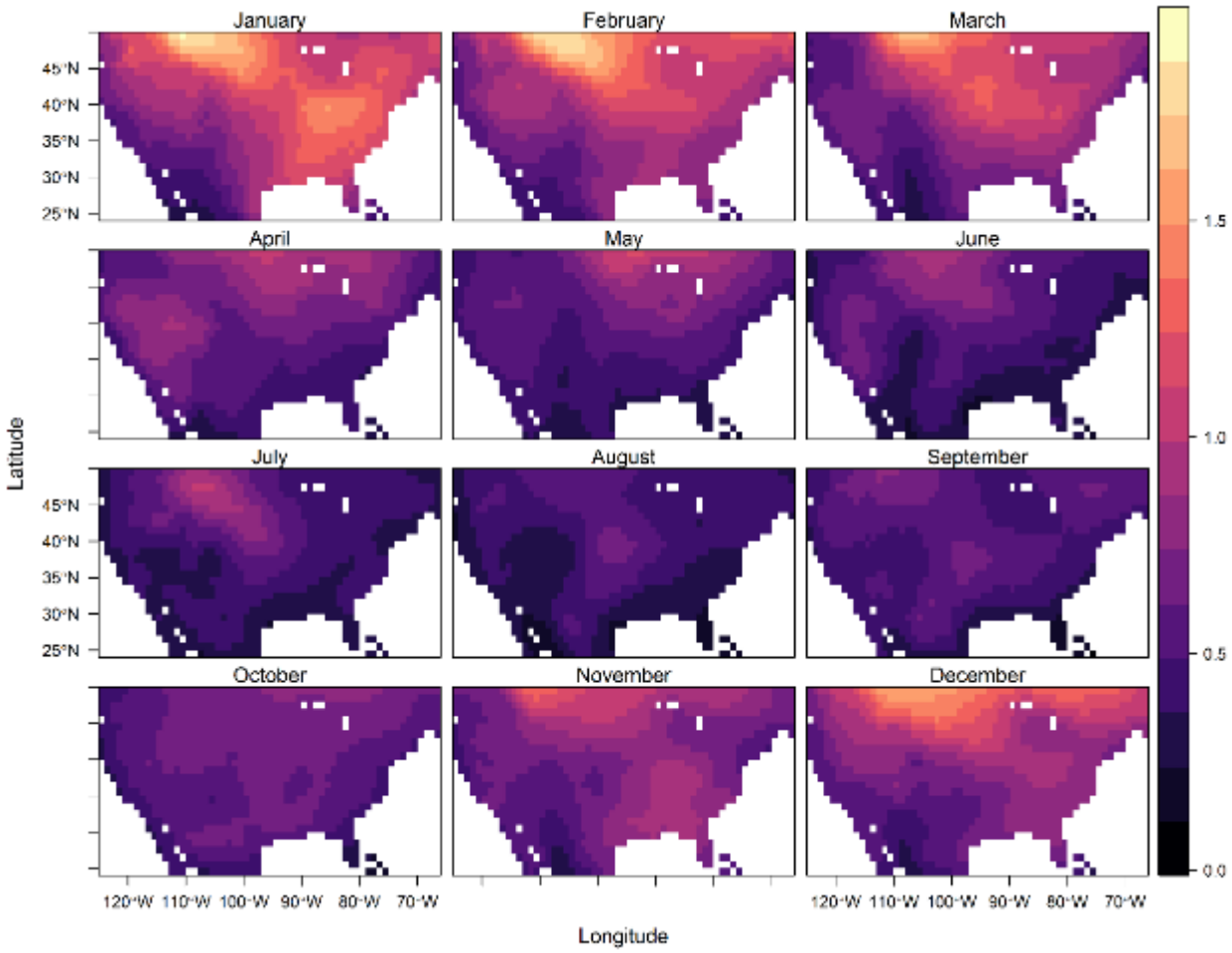

Figure 2-7 The variability induced by ICV across the CONUS as the standard deviation from the 1000-member OBS-CESM.

The variability (noise) induced by ICV are shown in Figures 2-7 to 2-9, as a result from the OBS-CESM and OBS-MACA, respectively. Both observational large ensembles showed similar picture in terms of spatial and temporal distribution of the variability in the temperature trend: variability tends to be higher in the northern areas (cooler regions) than the southern regions, and winter months tend to have more variability than summer months. The noise estimated from the OBS-MACA tend to be higher than that of the OBS-CESM, as the downscaled datasets reveal more fluctuations in temperature over the years than the coarser dataset.

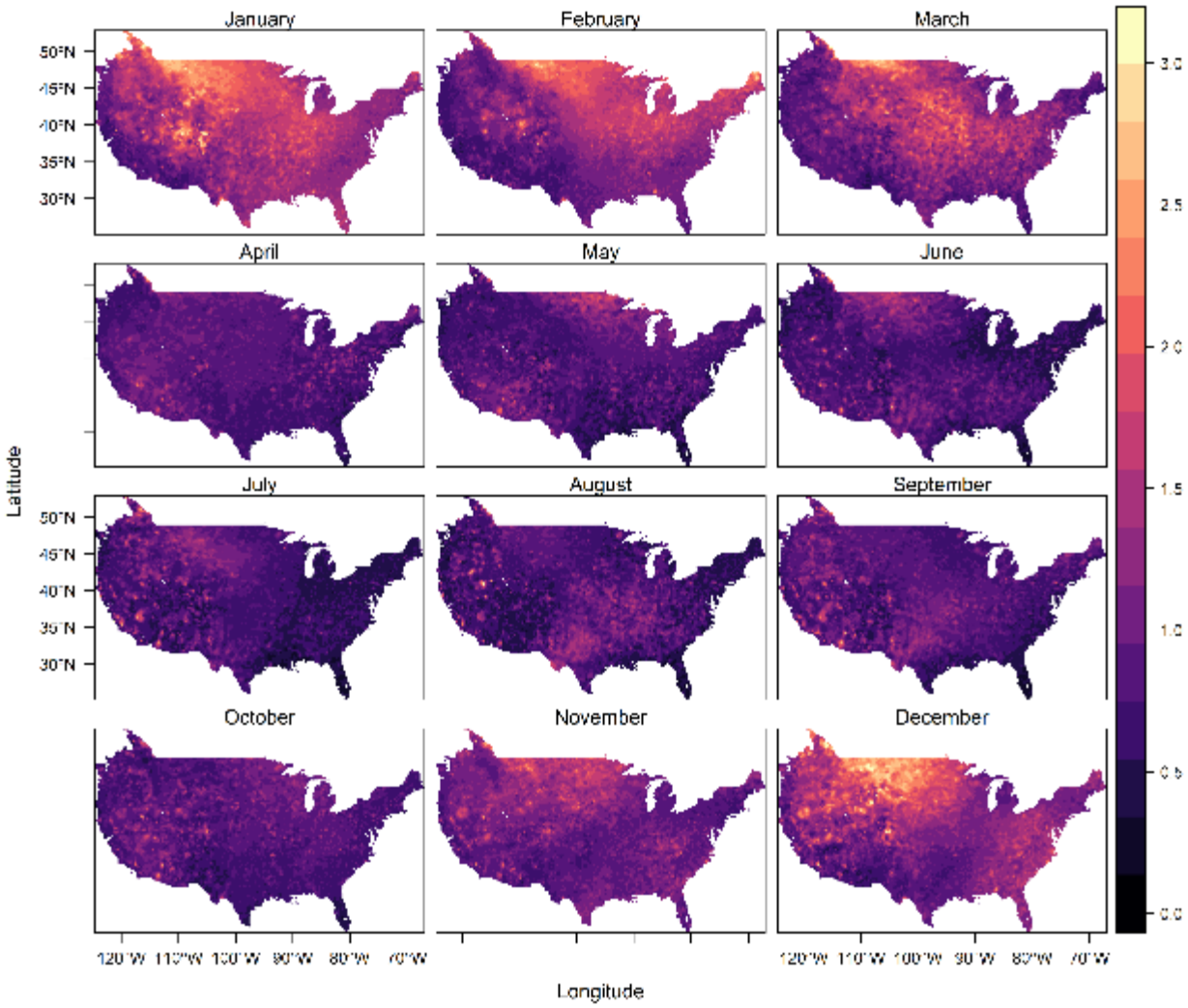

Figure 2-8 The variability induced by ICV across the CONUS as the standard deviation from the 1000-member OBS-MACA

Over the past period (1960-2010), mean temperature across the CONUS increased between 0-1.5ºC (on average 0.5ºC), with summer months and southwestern areas tend to increasing faster (see Figure 2-10); the southeastern region experienced slower temperature rise (between 0-1ºC). In the future period (2011-2060), temperature tend to rise much faster by mid-century according to the RCP 8.5 scenario, ranging from 1-5ºC (on average 1.27ºC) across the CONUS, with the summer time average increase of 1.52ºC (see Figure 2-10). January will see the highest rise in temperature, especially in the northeastern region. The southeastern region will experience an average increase of temperature 1.14ºC, with stronger increase in summer months by 1.27ºC (April to September).

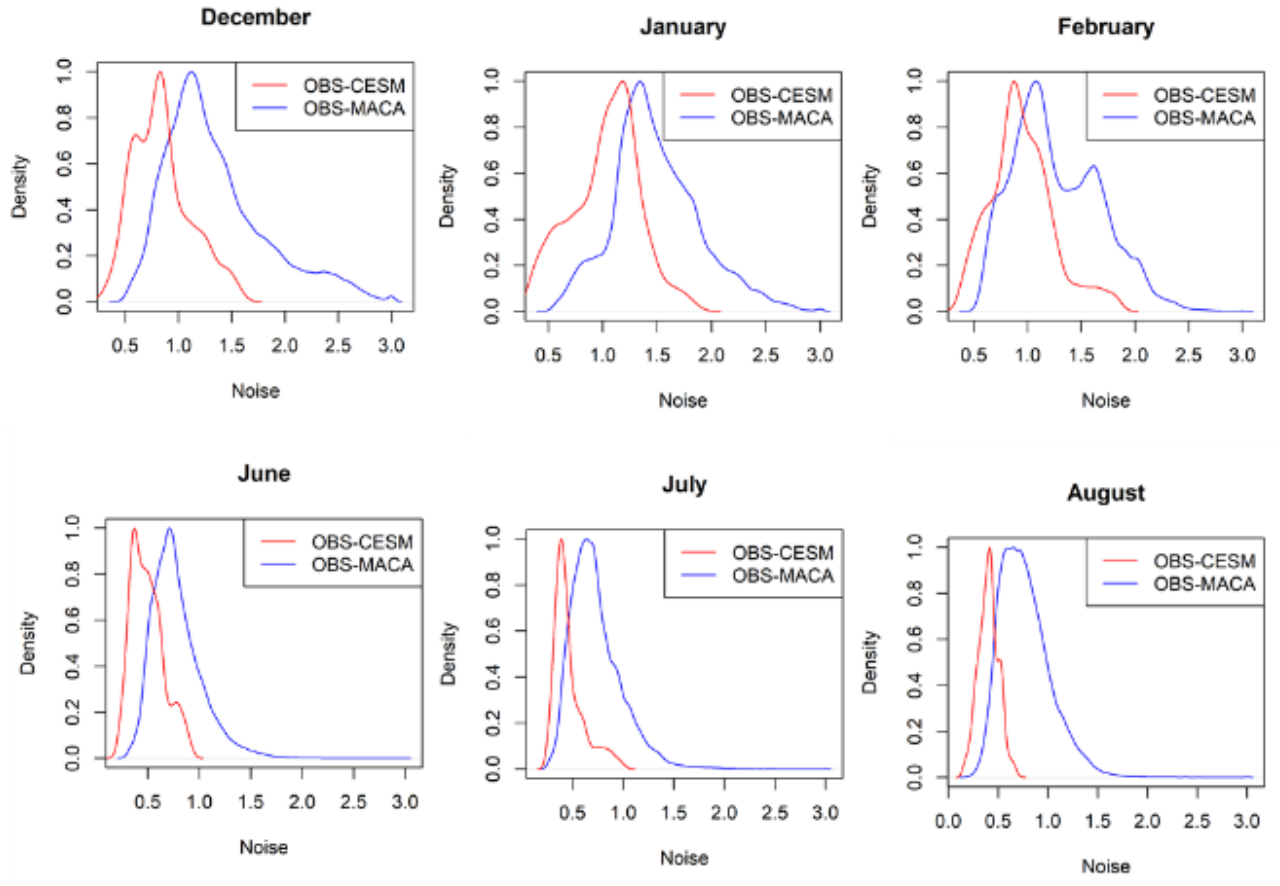

Figure 2-9 Histogram of noise (or variability) from two OBSLE show that variability increases with the downscaled data.

2.3.3 Signal-to-noise ratio of mean temperature trend

The signal-to-noise ratio map (Figure 2-11) reveals whether the temperature trend is significant or not. For the past period, the winter time for most of the CONUS did not show a strong signal (SNR < 2), except for the southwestern (SNR between 2-4 for October, November and December). The summer months for the southwestern and some parts of the southeastern area, on the other hand, had high SNR, indicating the rising trend in temperature during this season in those areas were significant. In particular, June, July and August - the three critical months during the growing season for maize and soybean - show strong SNR (4-6), indicating there has been a significant increase in mean temperature for those months during the historical period. This summer time rising temperature trend seems to continue in the future period (2011-2060), as indicated by high SNRs for the most part in the CONUS. The rising trend in winter temperature also tends to become more prominent, especially for the western U.S.

SNR estimates by the OBS-MACA ensemble showed that rising trend in temperature under the RCP 8.5 will be significant in the mid-century. While higher latitude regions may benefit from a warmer growing season, due to longer growing season and strong photosynthesis rate due to increase in radiation efficiency, lower latitude regions such as the southwestern and southeastern regions may face reduction in crop yield due to heat exposure, droughts, higher evaporative demands, etc. Outside of the growing season, warmer spring time (January, February and March), especially in the northern region, may lead to reduced snow pack, resulting in drought in the summer. This will pose a threat to water management and crop irrigation.

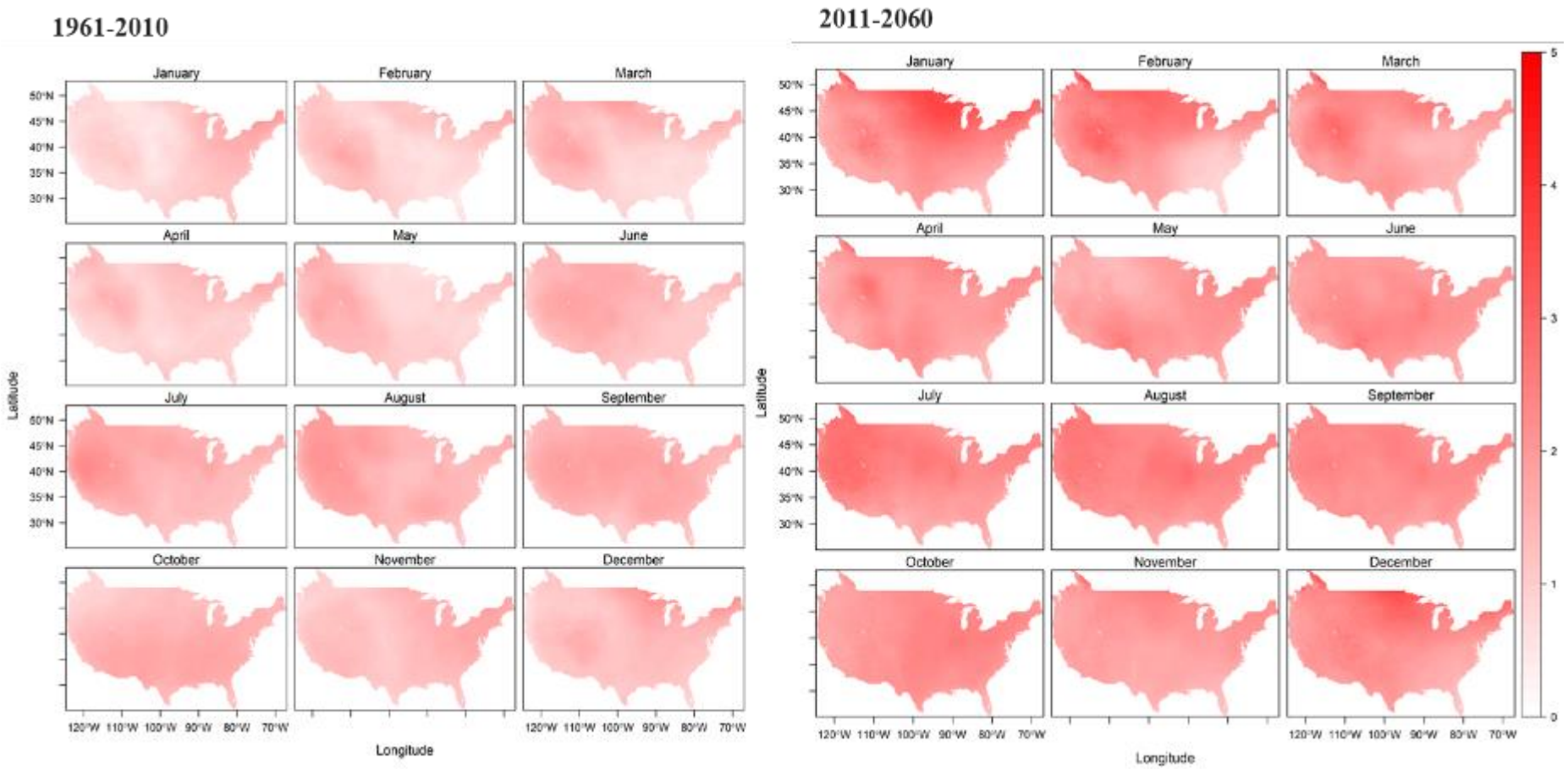

Figure 2-10 Signal (or magnitude) of mean temperature trend across the CONUS all months in the past (1961-2010) and the future (2011-2060) as a result from the high-resolution ensemble OBS-MACA

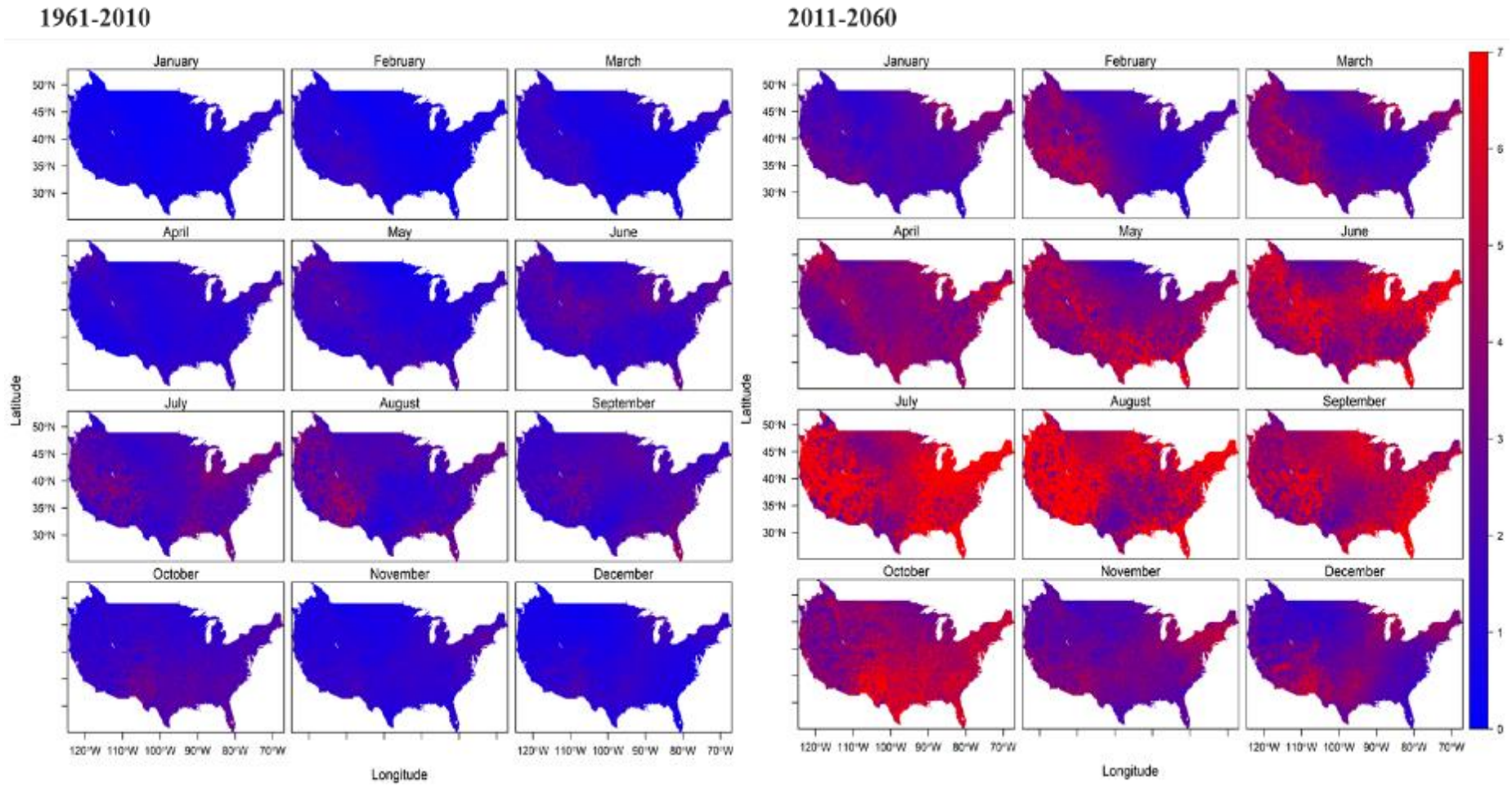

Figure 2-11 The signal-to-noise ratio for the mean temperature for the past (1960-2010) and the future (2011-2060). The signal is defined as the total increase in temperature during the period estimated by the MACA-L13 ensemble, and the noise is standard deviation in the 1000 observational ensemble

2.4 *Conclusions*

To improve the level of confidence in assessing impacts of climate change on the environment and socioeconomic sectors, internal climate variability (ICV) needs to be quantified and removed from climate change projections. By developing a statistical model using high resolution data to assess the strength of temperature trend in the CONUS, the study showed that that historically (1961-2010), CONUS-mean increase of temperature was approximately 0.5^{o}C, while under RCP 8.5 scenario, CONUS-mean temperature will increase by a 1.27^{o}C increase in the future (2011-2060). The SNR estimated from the OBS-MACA showed the historical increasing trend in mean temperature was only significant to some part in the CONUS (mostly southwester, some part of southeastern and northeastern region), with the critical months for maize and soybean production (June, July and August). However, the increasing trend will become more prominent in most part in the CONUS for the summer time (SNR between 4-6), where a CONUS-average temperature might increase by 1.52^{o}C. This poses risks of drought and heat exposure for crop production in lower latitude regions. Mean temperature in winter months for the western regions may also experience significant increase by over 1^{o}C, which might lead to decrease in snow pack in the spring.

Several limitations that have not been addressed in this study including: i) downscaling methodology for the GCM ensemble, and ii) the choice of model in the ensemble and climate change scenarios. For the former, this study will be expanded to other downscaling approaches, including other statistical and dynamical downscaling, to reduce the bias caused by the downscaling method. For the latter, as more GCMs are developed and advanced, and as knowledge of climate change scenarios are improved, more data can be used to develop new large-scale observational ensembles. The large observational ensemble based on high resolution data provides more details on the strength of signal in mean temperature change over the historical and future period, thus enables decision-makers to focus on areas and seasons where changes are significant. This is particularly important in long-term regional strategic planning for multiple sectors, including agriculture.

Chapter 3 Impacts of historical droughts on maize and soybean production in the southeastern U.S.

Abstract

Drought is a recurring natural phenomenon that has devastating effects on various economic, social sectors, as well as natural ecosystems and biodiversity. The alarming trend of yield losses to drought in the US agriculture is rising over the years, despite increasing financial investments and technological development to abate drought impacts. As various global circulation models project rising temperature and increased precipitation variability in the future, droughts are likely to be exacerbated in terms of frequency and severity, thus increasing the costs and damages in agriculture. The impacts of drought on crop production depend not only on severity, but also the timing of the drought event. This study aims to: i) quantify yield losses due to extreme drought events during the growing season, and ii) mapping the spatial distribution of crop production loss risks region over the Southeastern U.S. Analysis of yield and drought correlation was conducted for two major food crops – maize and soybean – by detrending yield from the USDA-NASS records and modified drought indices SPI and SPEI for counties across the Southeastern U.S. The mapping of crop production sensitivity was conducted by combining a quantity of yield loss due to drought and the risk of drought, level of irrigation utility and drought indices. The initial results show that there is a significant negative correlation between yield loss and droughts in most counties. Yield losses (40% on average for maize and 25% for soybean) are significant when abnormally dry conditions occur. Maize is more susceptible to drought than soybean. As about 75% of maize and 64% of soybean are rainfed in the region, the increase in drought under warmer climate will lead to more water needs for crops.

3.1 *Introduction*

Drought is a recurring natural phenomenon that has devastating effects on various economic, social sectors, as well as natural ecosystems and biodiversity. Drought usually occurs as a result of reduction in precipitation over an extended period of time compared to the normal condition; however, there is no single quantitative threshold of precipitation deficiency that leads to drought (Wilhite, 2000; Wilhite and Glantz, 1985). Concurrent events such as high temperatures, strong winds and low humidity can also greatly aggravate drought severity. The consequences of droughts are heavy in terms of human life, economic damage and biodiversity loss. In the United States, NOAA (2020) estimated that from 1980 to 2019 drought has cost an

average of $252.7 billion (Consumer Price Index (CPI – adjusted)) in damages to infrastructure and agricultural losses - constituting 14.2% of the total damages by natural disasters - and killed 3,865 people.

Amongst all economic sectors, agriculture is particularly vulnerable to the impacts of drought. It is estimated that the average annual crop losses associated with drought in the Great Plains area of the United States were about in US$700 million in the 1975 (White and Haas, 1975). In 1995, the US Federal Emergency Management Agency (FEMA) estimated annual losses attributable to drought at US$6-8 billion (Wilhite, 2000), whereas in 2018, this number was estimated at US$10-14 billion (Yusuke, 2019). This reveals the alarming trend of increasing vulnerability to drought in US agriculture, despite increasing financial investments and technological development to abate drought impacts. As various global circulation models (GCMs) project a warmer climate and increased precipitation variability in the future, droughts are likely to be exacerbated in terms of frequency and severity, thus increasing the costs and damages in agriculture (Pachauri et al., 2014).

Drought is commonly categorized as one of four types: meteorological, hydrological, agricultural and socioeconomic drought (Wilhite and Glantz, 1985). Agricultural drought is typically the result of precipitation deficiency during the growing season of crops or livestock, coupled with possible increase in temperature, low humidity and strong wind that result in high evapotranspiration. It is further characterized by reduced soil moisture, diminished surface and underground water supplies, and plant and livestock phenology. The impacts of drought on agriculture, in particularly crop production, depend not only on severity, but also the timing of the drought event. Droughts occur at the beginning of the growing season may not affect the crop growth, but if drought stress starts before or during the development stage, it highly likely causes crop yield loss and/or even crop failure. During the development stage, crops requirement for water, nutrient, and CO_2 increases, leading to higher rate of ET. Drought stresses during this period usually result in disrupted crop growth processes, shortened growth stages to reach early maturity before senescence, reduced plant organ sizes for efficient use of resources and lowered biomass accumulation and grain yield formation (Tuteja and Singh, 2012; Ihsan et al., 2016; Hasanuzzaman et al., 2013). Compared to other weather hazards such as extreme heat, drought events usually last longer and reduce water supplies, thus leaving the agricultural production system even more

vulnerable (Lesk et al., 2016). Since drought is a constant feature of almost every climate regime, it is crucial to provide farmers and policy makers, especially in dominantly rainfed regions, information regarding the severity and timing of drought so that they can effectively develop strategies to adapt and alleviate the damages of drought to their crops. Especially, with the imminent threats of climate change, providing robust, reliable projections of drought severity and timing in the future will critical in developing actionable plans.

Though precipitation is proven to be the main driver of drought, studies have shown that temperature rise markedly affects the severity of droughts (Carrão et al., 2016; Ebrahimpour et al., 2015; Golian et al., 2019; Gray et al., 2016). As seen in the Chapter 2, there has been a general temperature increase during the last 60 years in the U.S, and a strong probability of increase in temperature under high emission scenarios, it is important to incorporate the impact of temperature on droughts. The southeastern U.S. has seen an increase in drought occurrences over the years (Melillo et al., 2014). However, unlike the Western region, drought impacts on crop production under the context of climate change has been less extensively studied, though this region has been home to large population and lucrative agricultural business. Melillo et al. (2014) states that the Southeast is experiencing increased temperatures and decreased water availability (exacerbated by population growth and land-use change). The projections from global circulation models point to an increase in temperature and increase in variability of inter-annual and intra-annual precipitation in the most of the southeastern states (Melillo et al., 2014). Rising temperature will lead to increase in evaporation; coupled with likely precipitation decrease, this will reduce the net amount of water supply for environmental, human and agricultural uses. Agricultural production in the Southeast is a large driver of economic activity (Economic Research Service, 2018). A majority of crop land in the area is rain-fed, according to the data by USDA National Agricultural Statistics Survey (NASS) (https://quickstats.nass.usda.gov/). Changes in the precipitation and evapotranspiration patterns may thus affect these crops and lead to increased irrigation needs (Elliott et al., 2014). The increased demand for irrigation will put additional strain in the water sector, while also intensifying the demand for electricity. These interactions may thus create multi-sectoral constraints.

Quantitative drought indices are usually used to assess drought and its impacts on crop production. Although there is no universal quantitative definition for this phenomenon (Yihdego et al., 2019), studies on drought and its impacts on agricultural production normally rely on

standardized indices based on observed climate data (Hao and Singh, 2015; Yihdego et al., 2019; Heim Jr, 2002; Lu et al., 2017). Traditionally, attempts to formulate a drought index usually involve different drought-related climate and/or physical indicators, such as precipitation, evapotranspiration (derived from temperature, solar radiation, relative humidity and wind speed), soil water availability, soil water balance, etc. into a numerical value that indicate the deviation of the condition from the long-term normal average (Hao and Singh, 2015; Yihdego et al., 2019). Earlier recommended drought indices usually incorporate a single aspect of drought, such as precipitation deficiency, difference of precipitation and evapotranspiration, or soil moisture deficiency (Yihdego et al., 2019; Heim Jr, 2002). These indices are estimated based on precipitation (Standardized Precipitation Index (SPI), Munger's Index, Kincer's Index, Blumenstock's Index, Antecedent Precipitation Index, precipitation and temperature (Standardized Precipitation Evapotranspiration Index (SPEI)), or soil water balance (Palmer Drought Severity Index (PDSI), Palmer Hydrological Drought Index (PHDI), Palmer Z-index, Modified Palmer Drought Severity Index (PMDI)) (Heim Jr, 2002). In recent years, more complicated, multivariate drought indices have been developed that combine multiple metrics to capture different aspects of drought conditions for efficient monitoring and early warning, usually through combining objective and subjective indicators, extracting data from water balance model, and developing multivariate statistical relationships between indicators. Drought indices based on vegetation processed from satellite image processing are also developed, such as Normalized Difference Vegetation Index (NDVI) (Peters et al., 1991), Vegetation Condition Index (VCI) (Kogan, 1995), Scaled Drought Condition Index (SDCI) (Rhee et al., 2010), Soil Moisture Agricultural Drought Index (SMADI) (Sánchez et al., 2016) and the Empirical Standardized Soil Moisture Index (ESSMI) (Carrão et al., 2016). The application of these indices, however, are usually limited by availability of data, especially over a large area, and computational capacity. To assess and monitor agricultural drought over a large-scale region in an effective and timely manner, a comprehensive, robust yet simple index is needed. Therefore, drought indices such as SPI (McKee et al., 1993), SPEI (Vicente-Serrano et al., 2010), PDSI, Palmer Z-index, PHDI and PMDI (Palmer, 1965, Heddinghaus and Sabol, 1991) are still widely popular thanks to their simplicity, standardized calculation method and ease in interpretation and comparison (Zipper et al., 2016; Leng and Hall, 2019). In 2009, SPI has been recommended by the World Meteorological Organization (WMO) to be the standard drought metric.

Precipitation-based drought indices such as SPI rely on two assumptions: i) the variability of other variables, such as temperature and reference evapotranspiration (ET_o), is negligible compared to that of precipitation, and ii) the other variables are stationary. This means, the importance of other variable is negligible, and droughts are controlled mainly by temporal variability. However, PDSI does not have the multi-scalar characteristic essential for both assessing different drought types (meteorological, agricultural, etc.). On the contrary, SPEI combines the sensitivity of PDSI to changes in evaporation demand (caused by temperature, relative humidity and wind speed fluctuations and trends) with the simplicity of calculation and the multitemporal nature of the SPI. Based on the original SPI calculation procedure, SPEI uses monthly difference between precipitation and ETo (referred to as the climatic water balance) that is calculated at different time scales.

SPI and SPEI have enjoyed immense probability amongst the drought research community, including agriculture, owing to its simple methodology and low data requirements and their efficiency in drought assessment (Tian et al., 2018; Sanchez-Lorenzo et al., 2012; Lorenzo-Lacruz et al., 2010; Lu et al., 2017). While there have been many studies focused on quantifying the impacts of droughts on crop yield in terms of yield loss in the U.S. and the world, as well as projecting changes in severity and timing of future droughts by using GCM data, there are a couple of gaps in these studies:

i) Many studies used SPI and SPEI to categorize drought events (see Table 3-1) and quantify yield loss as deviation from long-term yield average deterministically (reporting a fixed yield change) in response to droughts (Peña-Gallardo et al., 2019; Matiu et al., 2017; Leng and Hall, 2019; Zipper et al., 2016; Lu et al., 2017). For example, Zipper et al. (2016) used SPEI to examine meteorological drought sensitivity of US maize and soybean yields from 1958 to 2001 and found that drought negatively affected crop yield over most crop-producing areas in the U.S., and yield is most sensitive to short-term droughts (1–3 months) during the critical development stage (July-August). However, fewer studies have adopted probabilistic methods to measure the uncertainty in drought impacts towards crop yield (Leng and Hall, 2019), especially in the Southeastern U.S.

ii) SPI may not reflect the true drought condition for crops, as this index does not include components that are important to agriculture, such as soil moisture. Therefore, using the original SPI with precipitation data may result in a misleading assessment of drought for agricultural impacts. To address this, some studies modified SPI by using effective precipitation (EP) in place of the original precipitation, as the former is expressed as the portion of precipitation that is available for crop consumption (Tigkas et al., 2018; Zarei and Moghimi, 2019; Ebrahimpour et al., 2015). These modified indices have not been frequently used in the U.S. In addition, the estimation of effective rainfall in these studies is normally based on empirical equations that use only precipitation data and usually exclude soil data due to restrictions in soil data availability, especially over a large-scale region, leading to uncertainty in EP estimation and subsequently drought indices.

iii) Increase in atmospheric CO_2 levels have important effects on crop plants, including, among others, reduction in stomatal conductance and transpiration per unit of leaf area. Decline in stomatal conductance in various plants due to elevated atmospheric CO_2 levels results in a reduction of transpiration (Allen Jr, 1990; Idso et al., 1994; Morison, 1985; Norby and Zak, 2011). Some studies suggest that CO_2-induced effects on reference evapotranspiration (ET_o) may offset the impacts of higher temperature (Ramírez et al., 1996). However, studying future changes in droughts and drought indices have largely ignored the impact of elevated carbon dioxide concentration in the atmosphere.

Therefore, this study aims to fill in these gaps by using high resolution historical climate (1979-2019), GCM projections, and soil moisture data to: i) compare SPEI and different modified SPIs using various EP estimation methods to determine the most suitable indices for the region, and ii) estimate the extent of yield losses to droughts over the historical period and the probability of yield loss to different drought severity levels.

Table 3-1 Classification of drought severity
(x: drought index value)

Category	Range
Extremely wet	$x \geq 2$
Very wet	$1.5 \leq x < 2$
Moderately wet	$1 \leq x < 1.5$
Normal	$-1 \leq x < 1$
Moderately dry	$-1.5 < x \leq -1$
Very dry	$-2 < x \leq -1.5$
Extremely dry	$x \leq -2$

3.2 *Methodology and data description*

3.2.1 Detrending crop yield data

The U.S. maize and soybean yield datasets exhibit an increasing, non-linear trend over 1979-2019 (Figure 3-1). This trend has been driven mainly by technological advances and management practices, while the year-to-year variability is mainly caused by fluctuations in climate factors. Assessment of drought impacts on crop production requires removing the non-climate induced trend from the original data, or detrending. Detrending yield consists of two steps: first, yield trend is simulated by fitting a trend model through the observational data; then, the trend is removed through a decomposition model, either additive or multiplicative.

3.2.1.1 Trend fitting for yield data

Trend fitting methods vary in their complexity and applicability. Simple linear regression (Quiring and Papakryiakou, 2003; Lu et al., 2017), second order polynomial linear regression (Trnka et al., 2007; Goldblum, 2013; Lu et al., 2017), moving average model (Lu et al., 2017), locally weighted scatter plot smoothing (LOWESS), and smoothing spline are some of the more common detrending methods. Lu et al. (2017) quantitatively compared the effectiveness of these models in detrending crop yields at the county and state level in the U.S and found that LOWESS is more appropriate than other methods due to its flexibility and adaptability to the length of the data and the underlying pattern of the non-linear and non-stationary yield time series at both the

county level and state level. Thus, LOWESS is chosen as the detrending method for this study.

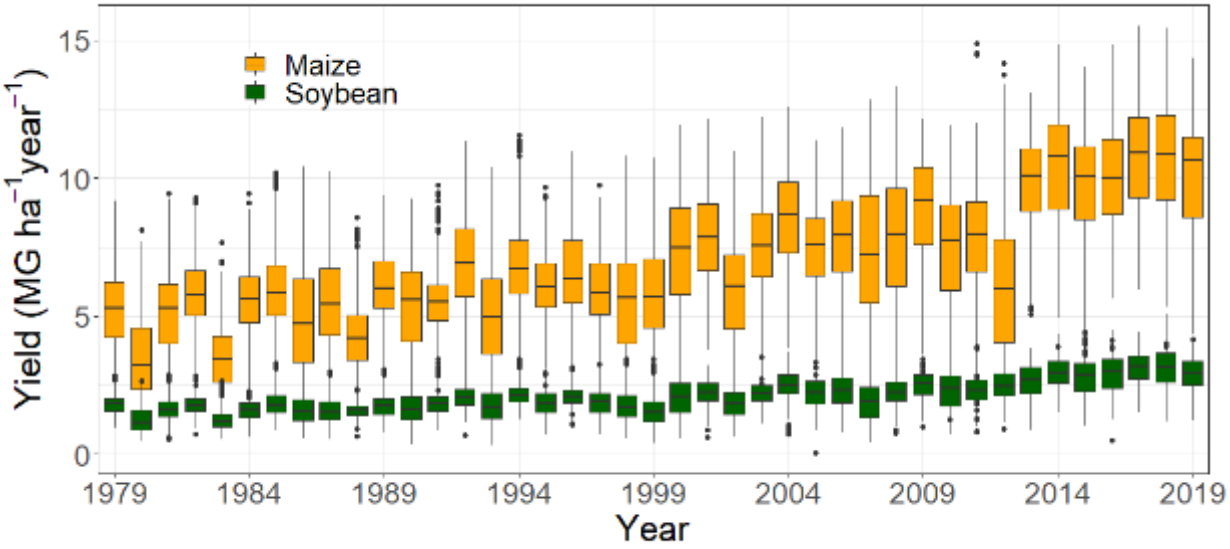

Figure 3-1 Historical yield trends of maize (yellow) and soybean (dark green) in the SERC region during 1979-2019. The box represents the 25th and 75th quantiles of county-level yields in the region, the black line represents the median yield, and the extended whiskers represent the 5th to 95th quantiles of yields for each year

LOWESS is a non-parametric smoothing regression model that involves fitting a weighted least squares model through a localized subset of data (Cleveland and Devlin, 1988; Cleveland, 1979). The weight of a point in the subset of data is inversely proportional to its distance to the chosen point of the interest in that subset, e.g.: the farther it is, the less weights it carries. In the dataset of size n, at every point i ($i = 1,2,\ldots,n$), a low-degree polynomial function is fit using weighted least squares to a subset of k data points near i. LOWESS is complete after regression values have been computed for each of the n data points. LOWESS requires calibration for two parameters: the fraction of data used in localized smoothing (f) and the degree of the polynomial weight function. The choice of these parameters will determine the performance of the LOWESS model in terms of robustness, effectiveness and efficiency. In this study, a quadratic function is chosen, as the lower order of polynomial will reduce the power of the model, while a third or higher order polynomial is too complicated and can lead to overfitting of data. The fraction f (to which the size of the subset data $k = n*f$), or sometimes referred to as the smoothing parameter, is usually selected through the cross validation. The smaller value of f, the more random error the trend may contain (overfitting), while a larger f value results in larger bias in detrending results.

The polynomial weight function is a tri-cube function, in which the weight of the neighboring point is determined by its distance from the point of interest (see equation (Eq.) 1).

$$\left.\begin{array}{ll} w(x) = (1 - |x|^3)^3 & (|x| < 1) \\ w(x) = 0 & (|x| \geq 1) \end{array}\right\} \quad \textit{Eq. 3-1}$$

where $|x| = \left|\frac{i-c}{dmax}\right|$ with $(i - c)$ being the distance between the point of interest *i* and the neighboring point *c*; and *dmax* is the maximum distance amongst the *k* neighboring points to the point of interest. Any points outside the subset of data ($|x| \geq 1$) will be given a weight of 0. For each point of interest *i*, weighted least squares with the weight *w(x)* of each neighboring point is estimated to minimize the error between the observation and the estimated point by the following formula:

$$\sum_{i=1}^{n} w(x) * \left(y_c - \beta_o^i - \beta_1^i * c - \beta_2^i * c^2\right)^2 \quad \textit{Eq. 3-2}$$

where y_c is the corresponding observation value at point *c*, β_o^i, β_1^i and β_2^i are coefficients of the fitted quadratic polynomial function at point *i* estimated by minimizing Eq. 2.

In this study, an R package called "fANCOVA", which contains a collection of R functions to perform nonparametric analysis of covariance for regression curves or surfaces was used to conduct the LOWESS fitting. The function "loess.as" in this package performs LOWESS fitting for the dataset, with automatic selection of the smoothing parameter using bias-corrected Akaike information criterion (AIC).

3.2.1.2 Removing trends from yield time series

After yield trend is fitted, the yield data series is detrended by removing the trend from the dataset, and the resulting residuals are detrended yield data. There are two common methods for removing the trend: i) additive method, in which the original value is subtracted from the trend line, and ii) multiplicative method, wherein the residual is computed as the ratio of the original value to the trend line value. The choice of the decomposition model depends on the variance of the observational data. For datasets with constant variance over time, the additive model is more appropriate since it assumes the trend and the residuals to be additive. The multiplicative method is more suitable to non-stationary datasets (statistical properties such as mean and variance are not

constant over time) for removing the trend in the variance of the original dataset that could be carried over to the detrended dataset.

One of the tests for stationarity of time series data is the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. This test aims to detect the existence of a unit root (a stochastic trend that causes unpredictable pattern in a time series). The KPSS test is based on linear regression, in which the time series x_t is broken into three components: a deterministic trend (β_t), a random error (r_t) and a constant error (ε) for each time step t.

$$x_t = r_t + \beta * t + \varepsilon \qquad \textit{Eq. 3-3}$$

If the time series is stationary, r_t will be the same for every data point. The test uses ordinary least square to find equation (3) with x_t usually being log-transformed to comply with the normality assumption. The null hypothesis of the test is the data is stationary. If the test statistics is above a certain critical threshold for alpha levels (p-value) of 0.1, 0.05 and 0.01, then the null hypothesis is rejected at the level, and the data is non-stationary. To test the stationarity of the detrended time series by the additive and multiplicative decomposition methods, the KPSS test is applied to both time series at the county level. P-values exceed 0.1 indicates that the time series is stationary.

3.2.2 Effective precipitation estimation

The definition of effective precipitation (EP) depends on the field of application (hydrology, meteorology, etc.). In agriculture, EP is defined as the portion of precipitation that is stored in the rootzone to be used effectively by plants (Martin and Gilley, 1993). Various processes and factors are involved in determining the effectiveness of precipitation, including climate, soil properties, antecedent soil moisture, vegetation cover, runoff, groundwater levels, topography, etc. (Martin and Gilley, 1993). To properly measure all of these processes involved in quantification of EP is almost impossible, thus simplified methods have been developed to estimate EP. These methods are based on varying degrees of representation of the hydrological cycle and on the application of interest, such as project planning, field-level irrigation scheduling, or integrated water basin management, etc. In general, there are non-empirical (weighted lysimeters, soil water balance modeling) (Dakshinamurti et al., 1961; Ebrahimpour et al., 2015) and empirical methods (empirical equations developed based on observational data) (Ali and Mubarak, 2017; Martin and Gilley, 1993; Painter et al., 2015; Zarei and Moghimi, 2019). Non-empirical methodologies are normally appropriate for field level studies, as they require data that may not be available at this

scale of study or sufficiently long for drought identification. Empirical equations provide broad yet useful estimations of EP for the purpose of understanding and predicting the study area's drought impacts on crop yields. There are several equations developed at different spatial scales, most notably are the Food and Agriculture Organization of the United Nations (FAO) equation (Brouwer and Heibloem, 1986), CROPWAT model equation (Smith, 1992), the USDA-Soil Conservation Service (SCS) equation (USDA) and its simplified version (USDA-simplified) (Martin and Gilley, 1993) (see Table 3-2). All these methods require monthly total precipitation *P* (mm). These methods are selected due to their simplicity in data requirement and their widespread application in different climatic conditions (Adnan and Khan, 2009; Ali and Mubarak, 2017).

Table 3-2 Empirical formula for effective precipitation calculation

Method	Formula	
FAO	$EP = 0.6 * P - 10 \left(P < 70 \frac{mm}{month}\right)$ $EP = 0.8 * P - 0.25 \left(P \geq 70 \frac{mm}{month}\right)$	*Eq. 3-4*
CROPWAT	$EP = \frac{P(125 - 0.2 * P)}{125} \left(P \leq 250 \frac{mm}{month}\right)$ $EP = 0.1 * P + 125 \left(P > 250 \frac{mm}{month}\right)$	*Eq. 3-5*
USDA-SCS	$EP = 25.4 * SF * (0.04931 * P_t^{0.82416} - 0.11565) * (10^{0.000955 ET_c})$ where SF is the soil water storage factor. SF is computed as: $SF = 0.531747 + 0.295167*D - 0.05797*D^2 + 0.003804*D^3$ where D is the usable soil water storage, estimated as 40-60% of soil available storage water (AWS) in the root zone depth. In this study, D is estimated at 50% of AWS. ET_c: evapotranspiration for a specific crop (see section 3.4.1)	*Eq. 3-6* *Eq. 3-7*
USDA-SCS simplified	The simplified version of Eq. 5 is Eq. 7, when assuming that D is generally 75 mm (3 inches) - thus SF becomes 1, and ETc is 200 mm: $EP = SF * (1.9443 * P_t^{0.82416}) - 4.56$	*Eq. 3-8*

3.2.3 Reference and potential crop evapotranspiration

3.2.3.1 Estimation method

This study employs the FAO-56 two-step approach in estimating reference evapotranspiration (ET_o) and crop evapotranspiration (ET_c) (Allen et al., 1998). While the original SPEI uses the Thornwaite formula to calculate ETo due to its simplicity, in this study ETo is computed using the Penman-Monteith (PM) equation as this equation has been adopted by the FAO, the International commission on irrigation and Drainage (ICID), and the American Society of Civil Engineers (ASCE) as the standard procedure for computing ET_o and proved to be more accurate in various climatic conditions compared to the Thornwaite equation (Pereira et al., 2015).

The PM equation estimates evapotranspiration for a hypothetical grass at crop height of 0.12 m, fixed surface resistance of 70 s m^{-1} and an albedo of 0.23:

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma\left(\frac{C_n}{T} + 273\right)u_2(e_s - e_a)}{\Delta + \gamma(1 + C_d u_2)} \quad \textit{Eq. 3-9}$$

Where:

- ET_o: daily reference crop evapotranspiration (mm d^{-1})[6]
- R_n: daily net radiation at the crop surface (MJ m^{-2} d^{-1})
- G: daily soil heat flux density at the soil surface (MJ m^{-2} d^{-1})
- T: mean daily air temperature (^oC)
- u_2: mean daily wind speed at 2m height (m s^{-1})
- e_s: mean daily saturation vapor pressure at 1.5 – 2.5 m height (kPa)
- e_a: mean daily actual vapor pressure at 1.5 – 2.5 m height (kPa)
- Δ: slope of the saturation vapor pressure-temperature curve (kPa $^oC^{-1}$)
- γ: psychrometric constant (kPa $\circ C^{-1}$)
- C_n: numerator constant (by FAO definition at daily time step: 900 K mm $s^3 Mg^{-1} d^{-1}$)
- C_d: denominator constant (by FAO definition at daily time step: 0.34 s m^{-1})

Crop evapotranspiration is calculated as:

$$ET_c = ET_o * K_c \quad \textit{Eq. 3-10}$$

where K_c is crop characteristic, which is the ratio of crop evapotranspiration compared to reference evapotranspiration. As crop water requirements change during the growing season for different development stages, K_c values also change. FAO tabulated the lengths and Kc values for maize and soybean for four distinct growth stages: initial, development, mid- and late-season stages (see Table 3-3) (Brouwer and Heibloem, 1986). The planting date for maize and soybean in the Southeastern U.S. usually falls in April (maize) and May (soybean). In this study, planting date for maize is fixed at April 1st and May 1st for soybean.

Table 3-3 Crop characteristics (K_c) and length of growth stages for maize and soybean according to (Lu et al., 2017)

Crop		Initial	Development	Mid	Late	Total
Maize	Length (days from planting)	30	40	50	50	170
	K_c	0.3	(ln)(1)	1.12	0.35(2)	

[6] Reference crop evapotranspiration is defined as "the rate of evapotranspiration from an extensive surface of 8 to 15 cm tall, green grass cover of uniform height, actively growing, completely shading the ground and not short of water".

Soybean	Length (days from planting)	20	25	75	30	150
	K_c	0.4	(ln)	1.15	0.50	

(1) K_c for the development stage linearly increase from K_c for initial stage to K_c for mid-season stage

(2) Assuming that maize is harvested at low grain moisture

3.2.4 Drought index estimation

3.2.4.1 Original indices

McKee et al. (1993) defined that deficit in precipitation is connected to the deficits in usable water sources, including soil moisture, ground water, snowpack, streamflow and reservoir storage. Any impacts of drought associated with water demand exceeding water supply originate with one or more of these five usable supplies. The time period from the arrival of precipitation until water is available in each useable form differs greatly, thus the time scale over which precipitation deficits accumulate becomes significant to assessing and monitoring drought events. Agricultural droughts (soil moisture) typically have a much shorter time scale than hydrologic (groundwater, streamflow, reservoir) droughts. Various studies show that maize and soybean yields are more sensitive to short-term droughts (1 to 3 months) (Zipper et al., 2016; Lu et al., 2017). In this study, drought indices at the 3-month timescale is computed as this timescale is reasonable to account for immediate impact of drought on crop yields, and also take into consideration the cumulative moisture conditions of the previous months, which is important to crop growth. Drought indices at 3-month scale are shown to have strong correlation with crop yields (Lu et al., 2017; Zipper et al., 2016). To calculate drought indices at 3-month scale, precipitation data are aggregated for every three months, e.g., precipitation data for August is the cumulation of June, July and August.

SPI was introduced by McKee et al. (1993) to "serve as a versatile tool in drought monitoring and analysis." SPI became widely popular and was chosen by the WMO to be the standard drought index due to its simplicity and flexibility in time scale, which is significant to drought. This index standardizes the precipitation deficit, considering the time series of cumulative precipitation for various timescales (1-48 months). The time series is then fitted in a statistical distribution and transformed into normal distribution. Therefore, the average SPI value for the study area and the specific reference period is zero, and non-zero values indicate wet (positive values) and dry (negative values) periods. McKee et al. (1993) proposed the use of gamma distribution for SPI calculation with two parameters, while some other researchers have used the log-normal distribution with similar results (Tigkas et al., 2018). SPEI (Vicente-Serrano et al.,

2010) is similar to SPI, but it is produced by standardizing the difference between precipitation and evapotranspiration ($P\text{-}ET_o$), also referred to as climatic water balance. A probability distribution of the gamma family is also used for SPEI calculation, but with 3 parameters, as the values of the time series can be negative. Vicente-Serrano et al. (2010) proposed to use the log-logistic distribution as it "can account for negative values, and it is capable of adopting different shapes to model the frequencies of the [$P\text{-}ET_o$] series at different time scales." In this study, SPI and SPEI are computed using the R package "SPEI", using "gamma" distribution for SPI, and "log-logistic" distribution for SPEI.

3.2.4.2 Computing SPI with effective precipitation

The absence of important variables to drought (such as temperature, soil moisture, etc.) in SPI is a limiting factor for its application in agricultural drought assessment. However, since most empirical EP were developed based on observations of soil water balance, such limitation can be accounted for by using EP instead of the original precipitation data, resulting in more reliable estimation of SPI for agricultural drought characterization and enhances its relevance to assessing drought impacts on crop production (Tigkas et al., 2018). In this study, EP is calculated using the Eq. 3-7. These data are then used as input for computation of SPI values, in addition to the original SPI. The results are referred to as SPI_FAO (using the FAO equation), SPI _CROPWAT (using the CROPWAT equation), SPI_USDA_simplified (using the simplified USDA equation) and SPI _USDA_CORN/SPI _USDA_SOYBEAN (using the USDA equation for maize and soybean, respectively).

3.2.4.3 Comparing drought indices

Different drought indices during the historical period (SPEI, SPI, SPI_CROPWAT, SPI_FAO, SPI_USDA_simplified, SPI_USDA_CORN/SOYBEAN at a 3-month scale) are compared to each other on the basis of their correlation with maize and soybean yields at the county level for each month during the growing season. The Pearson correlation coefficient *r* is used, and the significance (*p-value*) of the correlation is tested using a two-tailed t-test. Then, for each drought index, scale, month and crop, the number of counties with significant strong correlation ($|r| > 0.5$ and *p-value* < 0.05) between crop yield and the drought index is estimated. Indices of a particular timescale and month that show the strongest correlation with the crop are chosen for the further analysis.

3.2.4.4 Assessing drought impacts on crop yields

In this study, drought impacts on crop yields are assessed spatially and temporally through a couple of statistical methods.

Crop sensitivity to drought: a linear regression is fitted through the additive detrended yield (ΔY_a) and the drought index (I) time series for each county.

$$\Delta Y_a^{c,t} = \beta^c + \alpha^c * I^{c,t} + \epsilon^{c,t} \quad \textit{Eq. 3-11}$$

Where *c* and *t* indicate county and year, *I* is the drought index, β is the intercept, α is the slope of the linear regression, and ϵ is the residual. The slope α measures the sensitivity of the crop yield to drought impact, and a positive slope suggests that yield loss increases with drought severity.

It is suspected that areas with high level of irrigation and or high precipitation are those that have less drought impacts. To examine this, the slope α of from the linear regression of each county is plotted against the precipitation and irrigation level of that county. Irrigation data is aggregated by using the Moderate Resolution Imaging Spectroradiometer (MODIS) Irrigated Agriculture Dataset for the United States (MIrAD-US) (Pervez and Brown, 2010).

Yield loss to drought: each county and year is classified into different categories of drought severity based on their drought index value (see Table 3-1). Yield difference between the yield of a county-year that falls into a category and the expected yield of that county-year reflects the change in yield as a result of the dry/wet condition. Here, the yield difference is expressed as the multiplicative detrended yield (ΔY_m), which is the percentage of yield deviating from the expected yield. To measure yield loss due to drought across the SERC, all county-year ΔY_m of a category is weighted-average, with the weight being the county crop area.

$$\overline{\Delta Y_m} = \Delta Y_m^c * w_c \quad \textit{Eq. 3-12}$$

Where w_c is the crop area weight. Crop areas for maize and soybean can be taken from https://www.nass.usda.gov/. To obtain the robust confidence level and account for possible spatial correlation between detrended yield data, a bootstrap method where years are randomly sampled 1000 times with replacement is used (Schlenker and Lobell, 2010).

Yield loss probability under droughts: a copula-based probabilistic model is developed. Copulas (Nelsen, 2007) are functions that couple multivariate joint distribution functions to the

marginal distribution of each variable. Two advantages of copulas are they allow modeling dependency between variables that do not follow the same distribution, including non-normal distribution, thus avoiding prior assumptions about the relationship between the variables, and they can properly treat the tails of distribution, which is important for extreme drought assessment (Leng and Hall, 2019). A joint probability distribution $F_{XY}(X,Y)$ between the drought index (x) and detrended crop yields (y) is fitted through the data using one of the bivariate copula families:

$$F_{XY}(X,Y) = C[F_X(X), F_Y(Y)] \quad \textit{Eq. 3-13}$$

Where Fx(X) and Fy(Y) are marginal distributions of x and y, respectively. C is the cumulative distribution of the copula. After the joint distribution is developed, the conditional density probability $f_{Y|X}(y|x)$ of crop yield under a given drought condition can be expressed as:

$$f_{Y|X}(y|x) = c\left[F_x(X), F_y(Y)\right] * f_Y(y) \quad \textit{Eq. 3-14}$$

Where c is the copula function, $f_Y(y)$ is the probability distribution function (PDF) of crop yield. The probability of crop yield falling below the expected level (Y<0) under a given drought severity (X=x) is $F_{Y|X}(Y < 0|X = x)$, which can be estimated as the area under $f_{Y|X}(y|x)$.

A copula function is fit through the ΔY_m and drought indices to estimate the probability of yield falling below the expected yield is estimated ($\Delta Y_m < 0$) for a given drought condition. In this study, the fitting of the copula function is done with the R package "copula" and "VC2copula" which enables the selection of the best fit distribution function, optimal parameters and estimation of conditional density distribution.

3.3 *Result and discussion*

3.3.1 Detrended yields

Yield time series from 1979-2019 for county-level maize and soybean were detrended, and the residuals were estimated as the difference between the observational yield and the fitted yield (additive model) or the percentage of the difference to the fitted yield (multiplicative model) (Figure 3-2). KPSS test result show that there is the detrended time series by both decomposition methods are stationary. The results show that 90% of the maize yield difference over this period ranged between –2.21 to 1.79 Mg ha^{-1}, while 90% of soybean yield fluctuated between -0.58 to 0.48 Mg ha^{-1}. These are equivalent to a -41.3% to 31.6% deviation from the normal yield in maize,

and a -33% to 27% in soybean (see Figure 3-2). Maximum yield loss for maize (-86%) occurred in 2012, while maximum soybean loss (-96%) occurred in 2005.

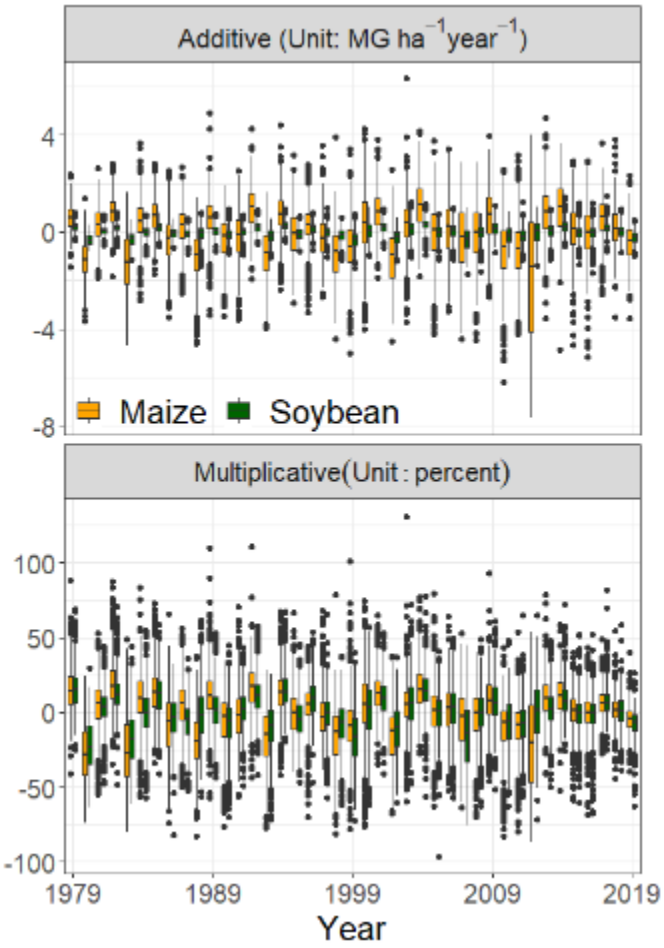

Figure 3-2: Detrended yield time series for maize (yellow) and soybean (dark green) over 1979-2019 with the additive (upper panel) and multiplicative (lower panel) decomposition methods. Each point represents a yield anomaly from the long-term yield trend estimated by LOWESS for a county. The boxplot represents the 25th-75th quantile range of yield anomalies across all counties for a year. The extended lines represent the 5th-95th quantiles. The highest and lowest points in each boxplot represent the maximum and minim yield anomaly for each year, respectively. KPSS test results for the detrended data by both decomposition methods show no change in variability over time, indicating that both methods result in stationary time series.

3.3.2 Historical drought conditions

Figure 3-3 shows the temporal pattern of droughts for each month of the growing season during the historical period (1979-2019) in the southeastern U.S. Drought indices are aggregated as the median value across the study region. The six indices show strong agreement in the overall drought pattern across the region, especially in April to June. However, in the month of July, August and September of 2012, SPEI indicates a higher level of drought condition compared to the SPI-based indices. The divergence can possibly be explained by the high temperature during those months in 2012, which leads to higher climatic evaporative demand (ET_o). The historical

period in this study covers some historical droughts, include 1980, 1988, 2002, 2011 and 2012. Spatial distribution of droughts in these years are shown in Figure 3-5.

3.3.3 Comparison of drought indices

Based on the estimated correlation coefficients r between drought indices and detrended yield data, SPEI and SPI_CROPWAT show superior performances as these indices have the highest percentages of counties that have strong correlation with both maize and soybean detrended yields (see Figure 3-4). In both crops and across all months in the growing season (April to September), the original SPI shows the least correlation with detrended crop yields. SPI_USDA_simplified tend to have stronger correlation with crop yields than the indices using the USDA equation for the specific crops. That SPI_USDA_CORN and SPI_USDA_SOYBEAN do not seem to improve from SPI_USDA_simplified can be attributed to the uncertainty in the estimation of crop-specific water requirement (ETc) over a large region, such as assumptions on the planting date (April 1st for maize, and May 1st for soybean in this study). Thus, Eq. (6) would be more accurate at a field scale with higher quality and location-specific data. Empirical, generic equations tend to be more useful in large-scale studies on the impacts of drought when there is a lack of data.

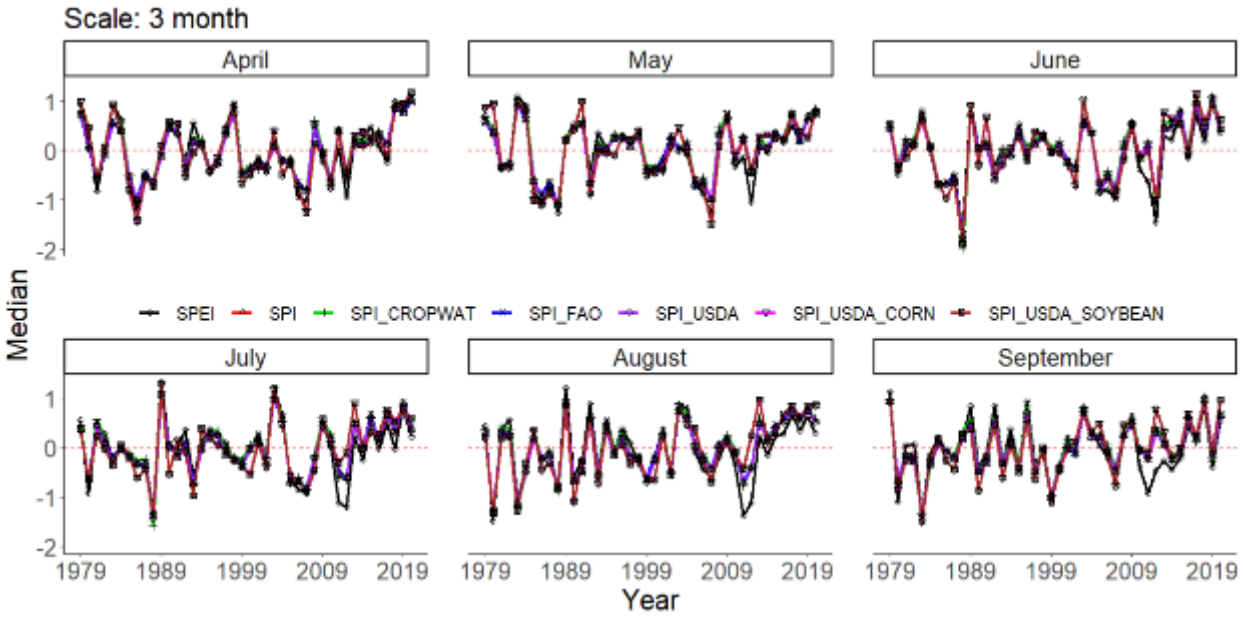

Figure 3-3 Growing season drought conditions during 1979-2019 based on six different drought indices on a 3-month scale. The x-axis is the value of the drought indices expressed as the median of all grid cells covering the southeastern U.S. SPI and SPEI

are estimated using the original precipitation, and the suffixes represent the formula used for estimating effective precipitation for the modified versions of SPI (see section 3.5.2)

Across all indices, drought in August (maize) and September (soybean) shows significantly higher correlation with crop yields than other months. The developmental (reproductive) stages of maize generally fall within July, whene silking starts. While soybean is generally grown later than maize (Sacks et al., 2010), the developmental stage of soybean starts earlier and lasts longer (roughly about three months) compared to maize, which explains the strong correlation with both August and September (Brouwer and Heibloem, 1986). Since the 3-month indices in August (September) are cumulative precipitation during the reproductive stage of maize and soybean, the strong correlation between drought indices in these months indicate the highly negative impacts of drought on crop yields. This finding is in agreement with previous studies that showcased the need for monitoring drought during this period (Lu et al., 2017; Zipper et al., 2016).

3.3.4 Drought sensitivity to crop yield

Based on results in section 4.3, SPEI and SPI_CROPWAT for August (maize) and September (soybean) are chosen for further analyses of drought impacts on crop yield. Figure 3-5 shows crop yield sensitivity to drought. Drought sensitivity is expressed as the slope α of the linear regression between detrended yield and drought indices (Eq. 37). SPEI and SPI_CROPWAT both show similar spatial pattern of drought impact on maize and soybean yield. Drought is mostly negatively correlated with crop production in the study region (see Figure 3-5). Some southernmost and coastal counties in the region exhibit negative relationships with drought. These counties have historically higher precipitation than the rest of the region (see Figure 1-2). Although the relationships found in these counties are mostly not significant at p-value <= 0.05, this finding suggests that areas that are historically wet would benefit from drier conditions.

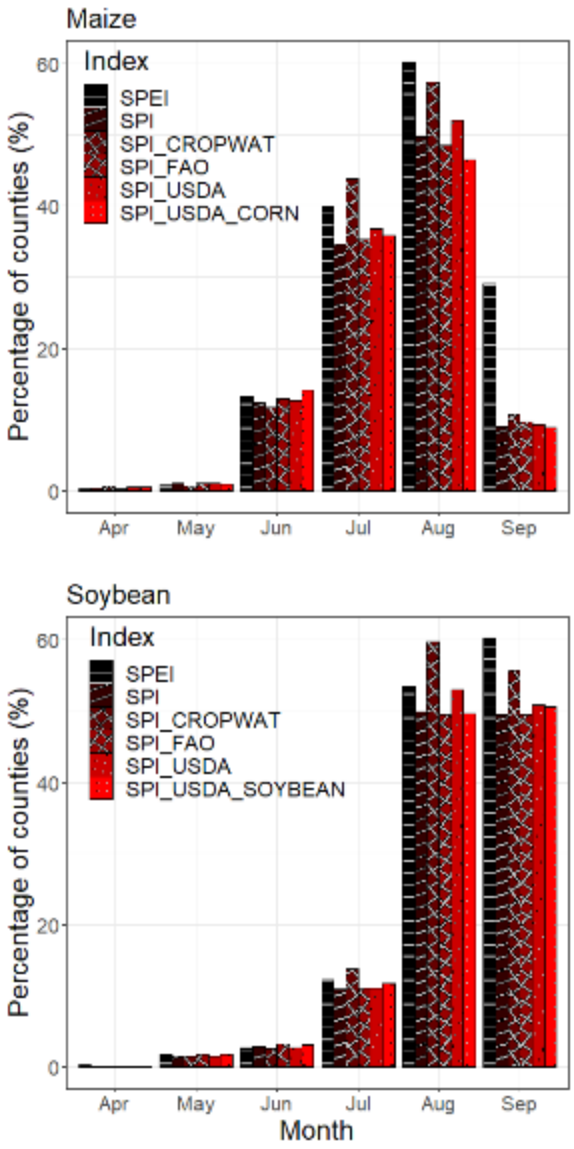

Figure 3-4 Performances of different drought indices expressed as the percentage of counties that exhibit strong correlation ($|r| > 0.5$ and p-value <= 0.05) with maize (upper) and soybean (lower) detrended yield

In general, drought has much greater negative impact on maize yield than soybean, which is reasonable given soybean has lower yield than maize. Increase in a unit of SPEI (SPI_CROPWAT) can reduce on average 0.62 (0.60) MG ha^{-1} of maize yield and 0.30 (0.28) MG ha^{-1} of soybean, and the range of sensitivity is (-0.45 – 1.61) for maize, and (-0.13 – 0.42) for soybean. On average, SPEI (SPI_CROPWAT) is responsible for 29.1% (28.3%) of year-to-year variability in maize yield and 30% (28%) in soybean yield, ranging from 0 to 80.1% (75.2%) for maize and from 0 to 85.3% (77.9%) soybean, as expressed by the adjusted R^2 values from the linear regression (see Figure 3-6). The average R^2 for maize and soybean in the southeastern U.S.

found in this study is about two

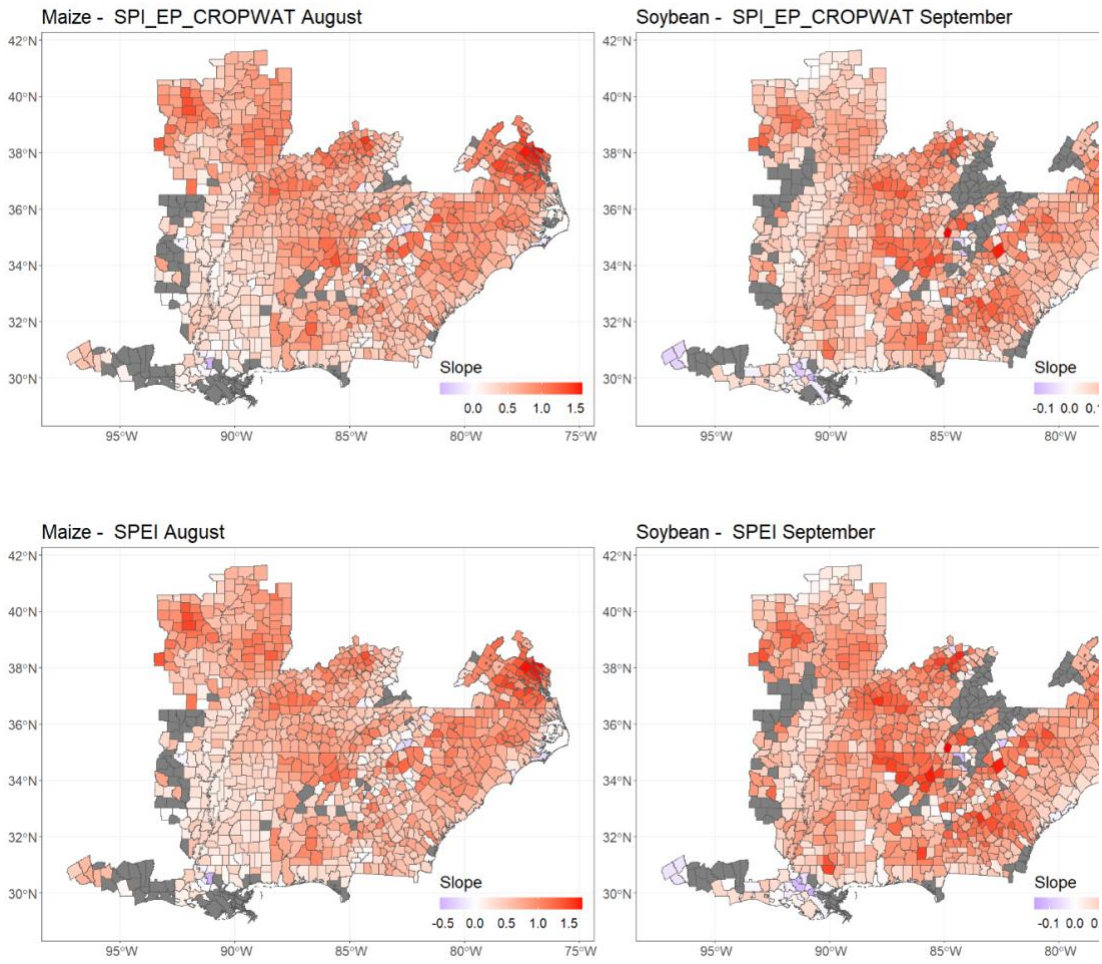

Figure 3-5 Maize (left panels) and soybean (right panels) sensitivity to drought, expressed as the slope of the linear regression between detrended yields and drought indices for two most critical months during the growing season. Grey

times higher than the national average 12.6% for maize and 12.8% for soybean reported by Zipper et al. (2016). This indicates that maize and soybean production in the southeastern U.S. is more sensitive to drought then the rest of the U.S., as maize and soybean in the region is mostly rainfed. That drought sensitivity in the southeastern is higher than northern maize and soybean producing regions in the U.S. (such as Nebraska and the Dakotas, northwestern Iowa, and eastern Wisconsin) as reported by Zipper et al. (2016) demonstrates that historically warm regions such as the south is more sensitive to increase in drought severity, and thus would be more vulnerable to future warming.

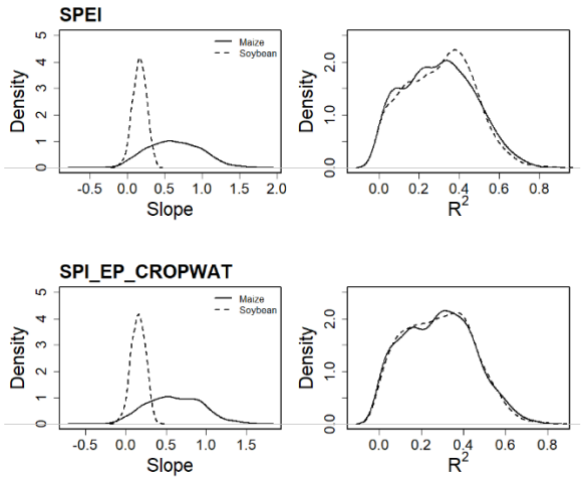

Figure 3-6 Distribution of crop yield sensitivity to drought, expressed as the slope of linear regression between detrended yield and drought indices (right panels) and year-to-year yield variation that can be explained by drought, expressed as coefficient of determination R^2 (left panel)

3.3.5 Yield loss due to drought

Figure 3-7 and Table 3-4 shows maize and soybean yield response to drought conditions in the developmental stage. Yield response is expressed as percentage of detrended yield to the long-term expected yield calculated by LOWESS. The region-wide yield response to drought severity is a weighted average from all the county-year detrended yields that experienced the drought condition. Yield response to drought conditions defined by SPEI and SPI_CROPWAT both show the significantly high level of yield loss under dry conditions, with the largest reduction under the extremely dry condition for both crops. Under extremely dry conditions according to SPEI (SPI_CROPWAT), average yield loss is -42.7% (-31.9%) for maize and -25.4% (-23.4%) for soybean (see Table 3-4). As the condition moves towards wetter conditions, yield change starts to become positive, showcasing the beneficial effects of sufficient precipitation on crop production for both maize and soybean. Maize and soybean yield experience the highest increase under moderately wet condition defined by SPEI (SPI_CROPWAT), with 8.20% (8.38%) for maize and 9.11% (7.56%) for soybean. However, when the condition becomes excessively wet, the benefit of precipitation diminishes for both crops. Soybean experience significant, though lower, yield increase under very wet condition, while this positive effect is not significant for maize. Under extremely wet condition, maize experiences little increase in yield, while the results by SPEI and SPI_CROPWAT diverge for soybean. Extremely wet condition as defined by SPEI shows very little increase in soybean yield, while the same condition defined by SPI_CROPWAT lead to a more, though not significant, positive impact. The negative impact of drought and the diminishing

effect of excessive precipitation found in this study agree with other studies (Schauberger et al., 2017; Li et al., 2019).

A previous study by Li et al. (2019) studied the impacts of drought and excessive rainfall on maize yield across the U.S. and found that excessive rainfall also has devastating impacts on maize yield, especially in northern maize-growing areas. The findings in this study are similar to Li et al. (2019) in terms of drought impacts, but the effect of rainfall on crop yield is more uncertain. In general, crop yields in the southeastern U.S. would be more at risk for drought than excessive precipitation, and crop production would benefit from wetter conditions compared to drier conditions. The southeastern U.S. is relatively warm, with mean temperature during the growing season. Several studies showed that maize and soybean yields in the warmer regions in the U.S. are lower than those in the cooler regions, and would experience higher yield reduction due to increase in temperature in addition to drying conditions, because the mean temperature is closer to the optimal temperature for crop growth (Hatfield et al., 2011; Rosenzweig et al., 2014). These results suggest that increase in droughts due to warming climate would pose great risks to maize and soybean yields in the southeastern U.S.

Using SPEI to quantify maize yield response to the extremely dry condition leads to higher yield loss estimation than using SPI_CROPWAT. SPEI incorporate both precipitation, temperature, and vapor pressure deficit (via evapotranspiration), and thus reflect the interaction between temperature and precipitation to crop yield. Previous studies have established the importance of such interaction to maize yield. Li et al. (2019) found that extreme drought coupled with extreme heat resulted in much larger maize yield reduction than without extreme heat. Lobell et al. (2013) found the predominant effects of extreme degree days on maize yield through simulations by the process-based model Agricultural Production Systems Simulator (APSIM), while precipitation has weaker impacts. The study found the effect of increase in temperature would lead to more water stress than decrease in precipitation, as it is associated with increased vapor pressure deficit, which contributes to water stress. While SPI_CROPWAT considers the soil moisture aspect through effective precipitation, it largely ignores the temperature effect, thus underestimates maize yield loss caused by extreme drought events. On the other hand, soybean yield loss estimated by using SPEI and SPI_CROPWAT under the extremely dry condition show little differences. Lobell and Field (2007) reported that precipitation has much larger impacts on

soybean, while maize variability is driven by temperature. Leng and Hall (2019) used SPI to estimate yield sensitivity to drought of several crops and found year-to-year variability in soybean

Table 3-4 Estimation of maize and soybean yield change for the southeastern U.S. in response to different drought conditions defined by SPEI and SPI_CROPWAT during the 1979-2019 period. Yield change is estimated as the percent of difference between detrended yield and the long-term expected yield

	SPEI			SPI_CROPWAT		
	Mean	**Maximum**	**Minimum**	**Mean**	**Maximum**	**Minimum**
Maize						
Extremely dry	-41.73	-48.72	-26.72	-31.96	-40.76	-21.38
Moderately dry	-14.72	-22.34	-8.10	-15.90	-23.06	-8.72
Very dry	-27.76	-35.15	-18.06	-20.97	-29.69	-11.86
Normal	3.63	0.55	6.47	3.09	-0.33	6.07
Moderately wet	8.20	4.89	11.52	8.39	4.79	11.32
Very wet	5.17	0.44	10.50	3.68	-0.96	8.65
Extremely wet	1.18	-6.06	8.09	2.09	-6.04	9.48
Soybean						
Extremely dry	-25.39	-32.82	-15.03	-23.43	-28.62	-15.71
Very dry	-19.82	-24.59	-14.30	-17.06	-20.58	-13.16
Moderately dry	-11.86	-15.07	-8.64	-9.74	-12.92	-6.08
Normal	1.96	-0.28	4.07	1.62	-0.88	4.11
Moderately wet	9.11	5.82	12.24	7.56	4.43	11.02
Extremely wet	0.38	-7.64	5.28	5.13	-2.08	8.42
Very wet	4.21	1.25	7.27	4.95	2.06	7.77

had the highest correlation with SPI. The similar results between SPEI and SPI_CROPWAT shows that soybean is more sensitive to deficit in precipitation rather than heat stress caused by

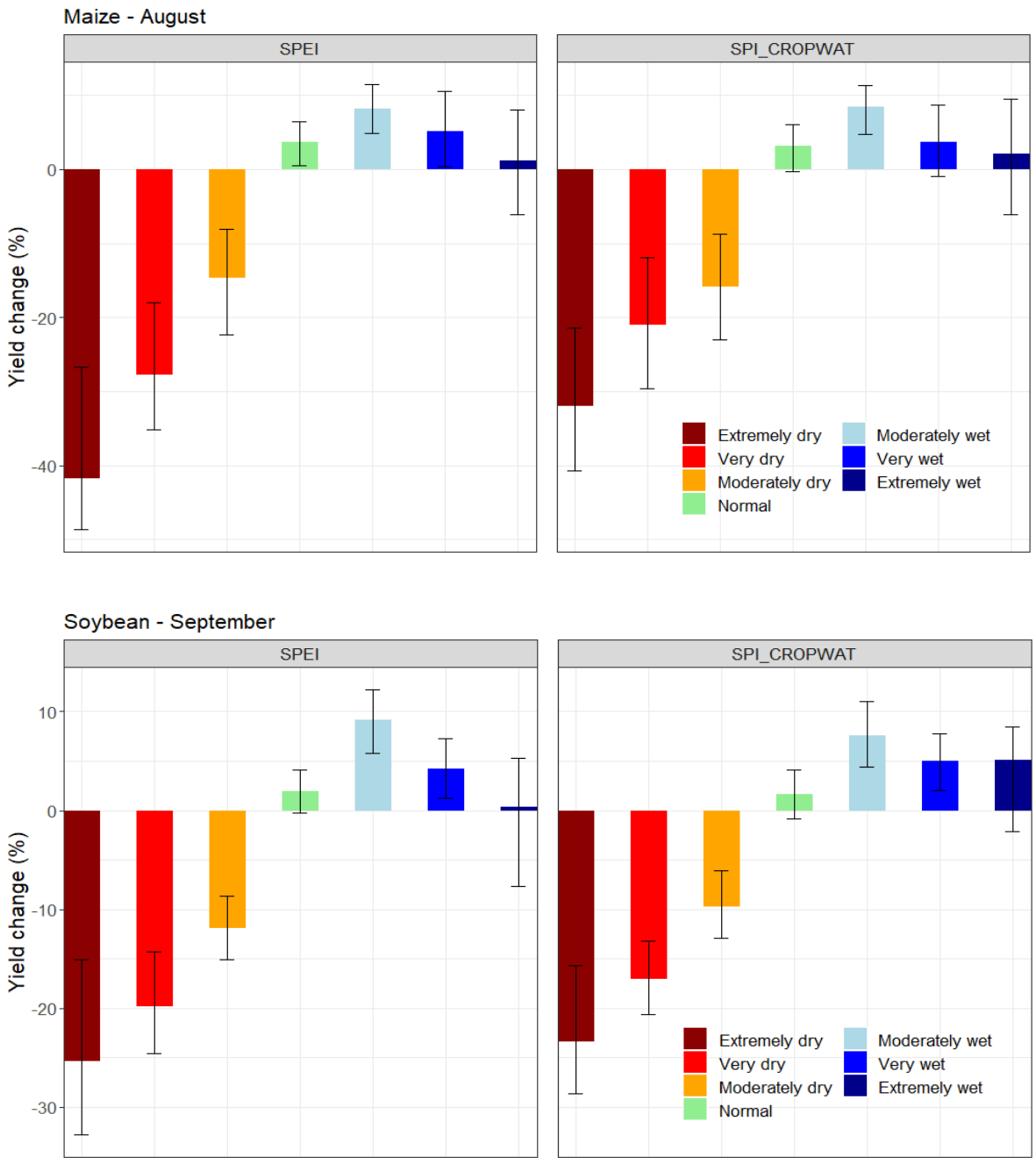

Figure 3-7 Averaged yield change in association with different wet/dry conditions in the critical months of the growing season for maize (upper panels) and soybean (lower panels). The bars represent percent of yield deviation from the expected yield for each wet/dry condition category, and the black whisker lines represent the 5- 95% confidence level from 1000-time bootstraps. Whisker lines not crossing 0 demonstrate that the impacts of the drought condition on crop yields are significant

temperature. It is suggested that climate change-induced temperature rise would pose a greater threat to maize, while changes in precipitation would be more responsible for yield loss in soybean in the future (Jin et al., 2017).

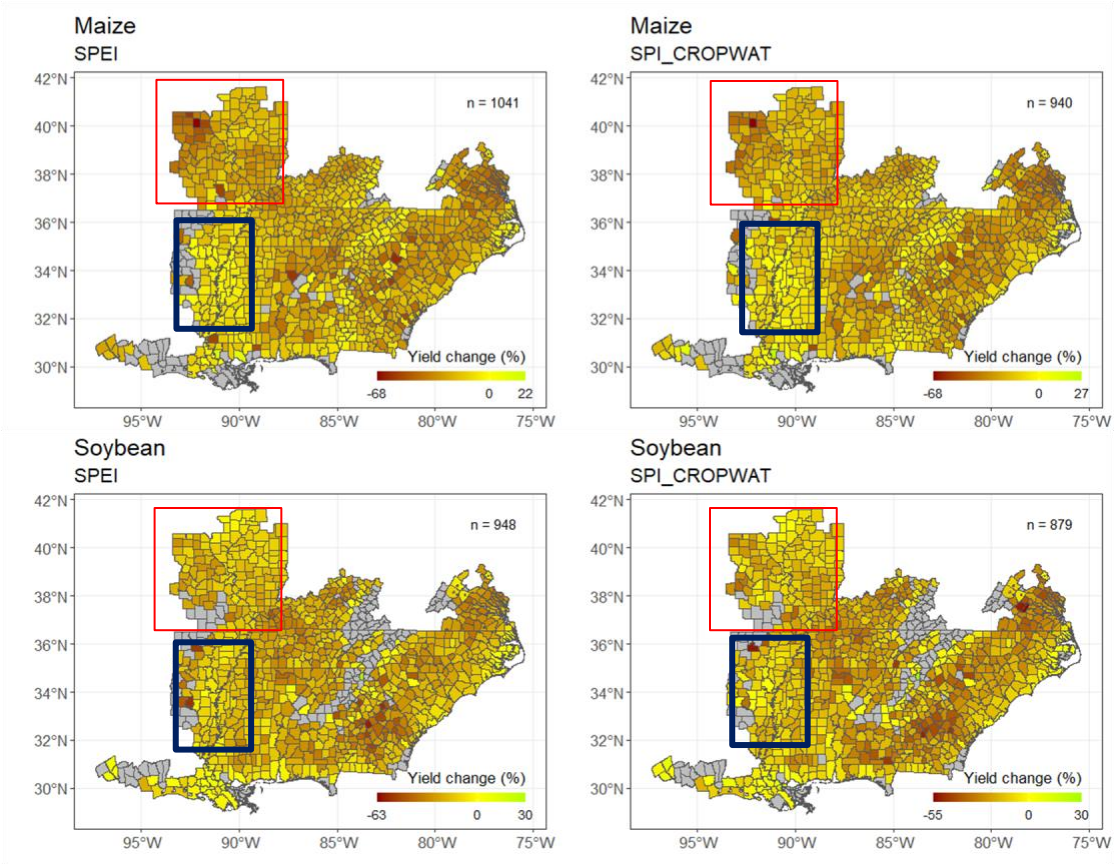

Figure 3-8 Spatial distribution of yield change (as percent of detrended yield to the expected yield) in response to droughts (drought index values ≤ -1) for all counties in the southeastern U.S. Yield change for each county is the average of the yield change of all the years when a drought occurs. The number of counties with yield data during the dry years is expressed as n*, and grey areas are counties that have no data. The red box indicates the high producing area of maize and soybean, and the blue box indicates the irrigated area*

Spatial distribution of yield loss due to the dry conditions (when index value ≤ -1) is shown in Figure 3-8. Using SPEI leads to more droughts being detected than SPI_CROPWAT. The high producing area (red boxes) shows high and moderate level of yield loss for maize and soybean, respectively, while the irrigated area shows smaller drought impacts on yield.

3.3.6 Probability of yield loss due to drought

By fitting a copula function through the detrended crop yield data and drought indices, the probabilities of yield change in response to different drought conditions were estimated. Figure 3-9 shows the probability density plot for yield change (as percentage of yield anomaly from the long-term expected yield) for a given drought condition in August (maize) and September (soybean), and Table 3-5 shows the probabilities of yield falling below 0 for different conditions

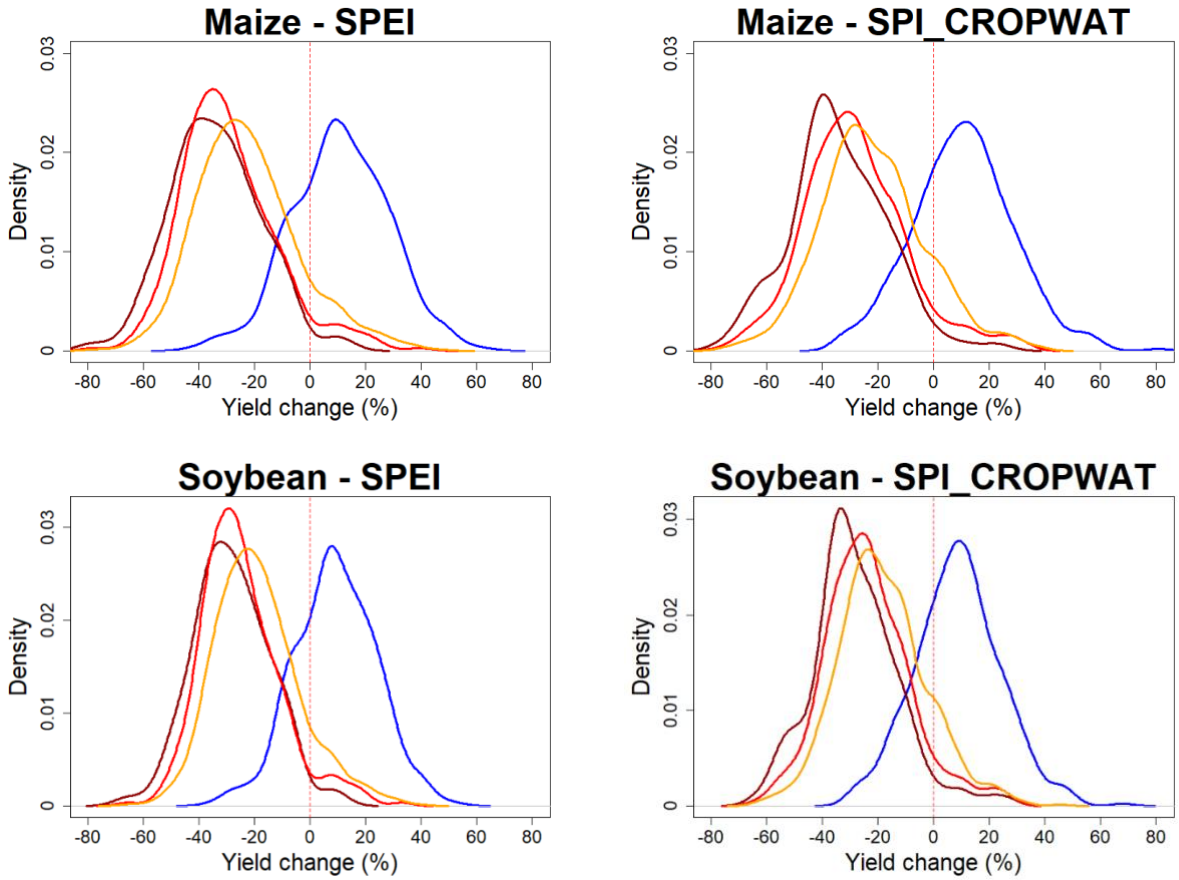

Figure 3-9 Probability distribution of yield loss in the southeastern U.S. for maize (upper panel) and soybean (lower panel) in response to different drought conditions defined by SPEI (right) and SPI_CROPWAT (left) as drought indices. The x-axis is percentage of detrended yield from the long-term expected yield, while the y axis is the density probability of yield change to a given drought condition. Drought conditions include moderately wet (blue), moderately dry (yellow), very dry (red) and extremely dry (dark red)

for maize and soybean. When drying condition occurs (drought index value is equal to or below -1), the probability of yield loss is more than 50%. As expected, both maize and soybean have the highest probability of yield reduction under the extremely dry condition, followed by the very dry and moderately dry conditions. Shifting from the normal dry conditions (-1 < drought index value ≤ 0) to the extremely dry conditions (drought index value ≥ -2) would increase the risk of yield reduction by over 40% for both crops (see Table 3-5). Even a shift from the normal dry condition

to moderately drought condition (-1.5 < drought index value ≤ -1) would increase the risk of yield loss by 11-12%. These results prove the high level of risk that crop production in the southeastern U.S. will face when drought events occur.

Table 3-5 Probability of yield falling below the expected yield for different drought conditions for maize and soybean

		Condition			
Crop	Index	Extremely Dry	Very Dry	Moderately Dry	Normal Dry
Maize	SPEI	86%	71%	56%	44%
	SPI_CROPWAT	84%	69%	55%	44%
Soybean	SPEI	86%	70%	55%	43%
	SPI_CROPWAT	83%	67%	54%	43%

3.3.7 Decoupling effects of irrigation on crop yield

The number of irrigated counties (where irrigation area exceeds 10% of the county area) in the southeastern U.S. is 74, with most of them are along the Mississippi Delta (see Figure 3-8). Among these counties, 71 counties grow maize and 72 grow soybean. It is suggested by other studies that irrigation can decouple the impact of drought on maize yield (Zipper et al., 2016). In this study, the differences in drought impact on maize and soybean between irrigated and non-irrigated counties were analyzed by analysis of variance (ANOVA) method (see Table 3-6). Maize grown in irrigated counties is significantly less sensitive to drought (p-value < 0.001). Soybean is also significantly less sensitive to drought in irrigated counties, though the difference is much smaller compared to that of maize.

Table 3-6 Differences between crop sensitivity to drought irrigated and non-irrigated regions. p-value *is based on the ANOVA of the slopes of the linear regression between detrended yield and drought index values.*

	SPEI			SPI_CROPWAT		
Crop	Non-irrigated	Irrigated	p-value	Non-irrigated	Irrigated	p-value
Maize	0.649	0.311	<0.001	0.625	0.311	<0.001
Soybean	0.166	0.131	0.0012	0.161	0.123	<0.001

3.4 *Conclusions*

Drought is one of the most devastating natural phenomena on agriculture. The impacts of drought on crop production depend not only on severity, but also the timing of the drought event. The study utilized high-resolution statistically downscaled data of 20 global circulation models

(GCMs) for two warming scenarios, RCP 4.5 and 8.5 to examine future conditions. Analysis of yield and drought correlation was conducted for two major food crops – maize and soybean – by detrending yield from the USDA-NASS records and modified drought indices SPI and SPEI for counties across the Southeastern U.S. Results show a significant association between droughts and yield reductions during the past decades. The mapping of crop production sensitivity was conducted by combining a quantity of yield loss due to drought and the risk of drought, level of irrigation utility and drought indices. The initial results show that there is a significant negative correlation between yield loss and droughts in most counties. Yield losses (40% on average for maize and 25% for soybean) are significant when abnormally dry conditions occur. Maize is more susceptible to drought than soybean. The increase in drought under warmer climate is expected to lead to more water needs for crops. Future projections show that it is likely that the frequency of extremely dry events will increase in the future under RCP 4.5 scenarios. Using modified drought indices (in which effective precipitation replaces original data) provides more meaningful relationship between crop production and droughts. This study hopes to provide visual information on the drought risk for informed decision making by stakeholders in the region.

Chapter 4 Crop net irrigation requirement

Abstract

Water supply plays a key role in crop production, as crops require sufficient amount of water to meet the water needs for optimal growth. This study employed a data-driven approach to estimate crop water demand under climate change impacts crop potential evapotranspiration (ET_c) and net irrigation requirement (NIR). Estimations from an ensemble of 18 downscaled GCM that show that there can be a wide range of projected changes in ET_c and NIR as compared to the historical baseline (1980 to 2009). Two scenarios of planting date were considered: the fixed planting date (crops are grown at the historical date) and adaptative planting date (planting date is chosen based on future climate) to examine the impacts of adaptation on NIR. The carbon fertilization effect (CFE) was also considered through two CO_2 level scenarios: the *contemporary* CO_2 level scenario (378 ppm) and the *elevated* CO_2 level scenario (where future CO2 levels change under RCP 4.5 and 8.5). The study finds that NIR tends to decrease under the fixed planting date scenario compared to the historical baseline due to a shortened growing season, and increase under the adaptive planting date scenario as the climatic demand for evapotranspiration increases due to higher temperature while the risk of accelerated maturity is reduced. The effects of elevated CO_2 level may reduce crop NIR under the same planting date scenario and RCP. Under the same planting date scenario and RCP, if CO_2 level is fixed at the contemporary level, NIR of maize will increase by 5.82 and 3.33 % in the near future and 4.18 and 1.37 % under RCP 4.5 and RCP 8.5, respectively, while NIR of soybean will increase by 13.53 and 4.51 % in the near future, and 9.10 and 2.05 % in the far future under RCP 4.5 and RCP 8.5, respectively. Spatial pattern of NIR shows that the southern area may experience little change or reduction in NIR due to increase in crop water demand under all planting date or NIR scenario as the impacts of increased temperature on shortened growing season length may overcome or cancel out enhanced climatic evaporative demand. The northwestern area near the Corn Belt, where crops were traditionally rainfed, may require investments in irrigation infrastructure, as this area may experience increase in NIR for both maize and soybean.

4.1 *Introduction*

Water supply plays a key role in crop production, as crops require sufficient amount of water to meet the needs crop transpiration and soil evaporation for optimal growth (also known as crop potential evapotranspiration - ET_c). Climate change is expected to increase water scarcity globally due to, to name a few, changes in precipitation patterns, warmer weather, increased vapor pressure deficit. (Vorosmarty et al., 2010; Schewe et al., 2014; Deryng et al., 2016). That means for rainfed crops, there might be a need for investment in irrigation if crop water demand cannot be met by precipitation; for irrigated crops, it means possibly less water availability to meet irrigation requirements. In addition, water managers and decision-makers, especially over large regions with competitive sectors, also have to face a water stress crisis as demands from other sectors, such as power generation and residential uses, rise amid declines in water availability in the future. To ensure sufficient supply of water to meet the demand of crops, there is a need to have a good estimations of crop evapotranspiration and crop net irrigation requirement.

Crop net irrigation requirement (NIR) is the difference between the demand (crop evapotranspiration) and the supply (water availability) of the water needed for crop growth. Crop evapotranspiration consists of two components: i) the climatic evaporative demand driven by climatic factors (temperature, relative humidity, wind speed, etc.), which is usually referred to as reference crop evapotranspiration (ET_o), and ii) the crop characteristic component, which is driven by crop type, crop phenology and physiology, such as growing season cycle, planting date, etc. The warmer weather predicted under most climate change scenarios may lead to higher overall ET_o. Studies have shown that ET_o is likely to increase under future climate due to warming. For example, Khalil (2013) estimated a ~12% increase of annual ET_o by the end of the century in Egypt. For the periods 2011–2040, 2041–2070, and 2071–2100 in the Xiangjiang River Basin, China, Tao et al. (2015) projected increases of 9.80%, 12.6%, and 15.6% respectively under Representative Concentration Pathway (RCP) 4.5, and 10.2%, 19.1%, and 27.3% respectively under RCP 8.5. Sunil et al. (2020) predicted an increase of ET_o by ~3% by the end of the century in the Upper Godavari basin (India). However, climatic changes also lead to changes in crop phenological characteristics, and as a result affecting overall ET_c. For example, an increase in temperature and subsequently growing degree days (GDDs) may shorten the growing season due to accelerated crop maturity, thus leading to lower total crop evapotranspiration during the growing season. Saadi et al. (2015) reported a 6.70% increase in annual ET_o for the period 2000–2050 for

wheat and tomato in the Mediterranean region, but reductions in ET_c of 6% (wheat) and 5% (tomato) by 2050 due to shorter growing season. Another study showed that under the RCP 8.5 scenario, total ET_c for several crops by the end of the century would be reduced: -5.10% (dry bean), -8.10% (corn), and -12.5% (spring wheat) (Scarpare et al., 2022). In addition, many studies have not included the alleviating effect of carbon dioxide on plant stomatal conductance, which reduces plant transpiration and increase crop water productivity (Vanuytrecht et al., 2012; Deryng et al., 2016; Taub, 2010). Another important aspect is adaptation to climate change via earlier planting, which also has an effect on ET_c. On average for the 2090s in the Columbia Basin Project area of the USA Pacific Northwest, the adaptative planting scenario (which allows for changes in planting date compared to the historical period) increases ET_c by 3.15% mm (dry bean) and 4.21% mm (corn), per season in the , while the fixed planting scenario (in which crops will be grown around their historical planting date) lead to a reduction in ET_c by 5.10% (dry bean) and 8.20% (corn) under RCP 8.5 (Scarpare et al., 2022).

Methods for the estimation of ET_c have been developed to meet the need for a robust and efficient estimation at different levels. Since its introduction in 1998, the two-step approach introduced in the FAO Irrigation and Drainage Paper No 56 "Crop Evapotranspiration" in 1998 (Allen et al., 1998) has been a standard approach to estimating crop evapotranspiration. First, ET_o is calculated by using the Penman-Monteith equation. This physically-based method to estimate evapotranspiration of a reference crop[7] has been well received by experts and proved to be applicable in a wide range of environments (Pereira et al., 2015). Then, ET_c is estimated by scaling ET_o with the crop characteristic coefficients (K_c). K_c values account for crop-specific influence on evapotranspiration. FAO standardized and tabulated the values of K_c to account for various crop growth stages and management conditions. This approach has been applied in many studies on crop water demand estimation and become the basis for various crop growth, soil water balance and hydrological models (Wane and Nagdeve, 2014; Lewis and Allen, 2017). It is probably the most widely applied method for estimation of crop water demand, development of water management models, and could be used for estimation of evapotranspiration both at the global and local scale (Acharjee et al., 2017b; Acharjee et al., 2017a; Bautista-Capetillo et al., 2013; Wriedt

[7] The reference crop by the FAO56 is "an extensive sur-face of 8 to 15 cm tall, green grass cover of uniform height, actively growing, completely shading the ground and not short of water"

et al., 2009). Although a number of field studies estimating ET_c for local conditions showed that FAO-56 overestimates total ET_c due to differences in FAO-tabulated K_c values and local values (Djaman et al., 2015; Djaman et al., 2018), the FAO-56 method is still highly applicable for large-scale estimations of ET_c due to lack of data available for local K_c and the resources, time and effort for collecting these data over a large area (Pereira et al., 2015). Other methods developed to estimate ET_c include energy or water balance methods. These include using lysimeters and a weather station at the site for direct measurement, process-based crop models with a soil water balance component (Anapalli et al., 2018; Martins et al., 2019a), or satellite-derived image of crop coverage (Bautista-Capetillo et al., 2013; Martínez-Cob, 2007; Spiliotopoulos and Loukas, 2019; Tasumi et al., 2006). However, these methods require a large amount of computational execution, resources and time for application over a large-scale region at which integrated water management is normally operated.

NIR requires estimation of the amount of the portion of precipitation that is stored in the rootzone to be used effectively by plants, also referred to as effective precipitation (EP) (Martin and Gilley, 1993). Various processes and factors are involved in determining the effectiveness of precipitation, including climate, soil properties, antecedent soil moisture, vegetation cover, runoff, groundwater levels, topography, etc. (Martin and Gilley, 1993). To properly measure all of these processes involved in quantification of EP is almost impossible, thus simplified methods have been developed to estimate EP. These methods are based on varying degrees of representation of the hydrological cycle and on the application of interest, such as project planning, field-level irrigation scheduling, or integrated water basin management, etc. In general, EP can be estimated by using water balance methods (such as weighted lysimeters, soil water balance modeling, etc.) (Dakshinamurti et al., 1961; Ebrahimpour et al., 2015), or mathematical formulas developed based on observational data over time and space (Ali and Mubarak, 2017; Martin and Gilley, 1993; Painter et al., 2015; Zarei and Moghimi, 2019). Water balance methods are normally appropriate for field level studies, as they require data that may not be available at the scale of this study or sufficiently long for drought identification. Mathematical equations provide broad yet useful estimations of EP for the purpose of understanding and predicting the study area's drought impacts on crop yields. There are several equations developed at different spatial scales, most notably are the Food and Agriculture Organization of the United Nations (FAO) equation (Brouwer and Heibloem, 1986), CROPWAT model equation (Smith, 1992), the USDA-Soil Conservation

Service (SCS) equation (USDA) and its simplified version (USDA-simplified) (Martin and Gilley, 1993) (see Table 2). All these methods require monthly total precipitation P (mm) and are popular in literature to their simplicity in data requirement and their widespread application in different climatic conditions. Amongst these methods, the USDA-SCS equation allows the consideration of soil characteristics (expressed via available water storage) of the location.

This study focused on estimating climate change impacts on NIR for the SERC region. This region is at risk of increasing droughts, and the regional crop production of may have to fiercely compete with other water-intensive sectors for water availability (Cheng et al., 2020b; Cheng et al., 2020a). Water managers in the region, especially around the Tennessee Authority Valley need to have robust estimations of potential changes in crop evapotranspiration and NIR to develop strategic plans. This study aimed to provide these estimations in a timely manner by applying the FAO-56 method to estimate ET_o and ET_c, and the SCS method for EP. The effect of elevated CO_2 concentration on stomal conductance, ET_c and eventually NIR was estimated by adjusting the bulk resistance coefficient of crops to new level of atmospheric CO_2. Potential adaptation by planting earlier was also considered by using a dynamic method to select suitable planting date in the future. This study projected ET_o, ET_c and NIR by using an ensemble of global climate models under two climate scenarios, RCP 4.5 and RCP 8.5 to consider the uncertainty in climate projections and emission scenarios.

4.2 *Methodology and Data*

4.2.1 Methodology

4.2.1.1 Crop evapotranspiration

4.2.1.1.1 Reference evapotranspiration ET_o

This study employed the FAO-56 two-step approach in estimating reference evapotranspiration (ET_o) and crop potential evapotranspiration (ET_c) (Allen et al., 1998). In the first step, the Penmann-Montieth (PM) equation was used to estimate ET_o. The PM equation has been adopted by the FAO, the International commission on irrigation and Drainage (ICID), and the American Society of Civil Engineers (ASCE) as the standard procedure for computing ET_o and proved to be more accurate in various climatic conditions (Pereira et al., 2015).

The PM equation estimates evapotranspiration for a hypothetical grass at crop height of 0.12 m, fixed surface resistance of 70 s m^{-1} and an albedo of 0.23:

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma\left(\frac{C_n}{T} + 273\right)u_2(e_s - e_a)}{\Delta + \gamma\left(1 + \frac{r_s}{r_a}\right)}$$ *Eq. 4-1*

Where:

- ET_o: daily reference crop evapotranspiration (mm d^{-1})[8]
- R_n: daily net radiation at the crop surface (MJ m^{-2} d^{-1})
- G: daily soil heat flux density at the soil surface (MJ m^{-2} d^{-1})
- T: mean daily air temperature (^oC)
- u_2: mean daily wind speed at 2m height (m s^{-1})
- e_s: mean daily saturation vapor pressure at 1.5 – 2.5 m height (kPa)
- e_a: mean daily actual vapor pressure at 1.5 – 2.5 m height (kPa)
- Δ: slope of the saturation vapor pressure-temperature curve (kPa $^oC^{-1}$)
- γ: psychrometric constant (kPa $\circ C^{-1}$)
- r_s: bulk surface resistance (s m^{-1})
- r_a: aerodynamic resistance (s m^{-1})

4.2.1.1.2 Crop evapotranspiration

Table 4-1 Crop characteristics (K_c) and length of growth stages for maize and soybean according to (Lu et al., 2017)

Crop			Initial	Development	Mid	Late
Maize	GDD [(i)]	Length [(i)]	30	70	120	170
		Stage	476	1112	1906 [(iii)]	2700
		Kc	0.3	(ln)[(ii)]	1.12 [(iv)]	0.35
Soybean	(days from planting)	Length	20	45	120	150
	GDD	Stage	327	735	1960	2450
		Kc	0.4	(ln)	1.15	0.50

i. Length of growth stage expressed in the number of days from planting date, while stage GDD is expressed as accumulated daily GDD from the planting date
ii. K_c for the development stage linearly increase from K_c for initial stage to K_c for mid-season stage
iii. Maize and soybean hybrid were assumed to be 2700- and 2450-GDD hybrid, respectively according to
iv. This is Kc value for maize harvested at low grain moisture

[8] Reference crop evapotranspiration is defined as "the rate of evapotranspiration from an extensive sur-face of 8 to 15 cm tall, green grass cover of uniform height, actively growing, completely shading the ground and not short of water"

Crop evapotranspiration is calculated as:

$$ET_c = ET_o * K_c \quad \textit{Eq. 4-2}$$

where K_c is crop characteristic, which is the ratio of crop evapotranspiration compared to reference evapotranspiration. As crop water requirements change during the growing season for different development stages, K_c values also change. FAO tabulated the lengths and Kc values for maize and soybean for four distinct growth stages: initial, development, mid- and late-season stages (see Table 3) (Brouwer and Heibloem, 1986).

The planting date for maize and soybean in the Southeastern U.S. usually falls in April (maize) and May (soybean). In this study, two scenarios for planting date are used: fixed calendar (in which planting date is fixed at April 15th and May 1st for soybean throughout the simulation period), and adaptive calendar (in which planting date is selected according to the temperature). The potential adaptive planting date for maize is chosen as the first date in the calendar year where air temperature is over 12.8^{o}C for 14 continuous days. This method for selecting potential planting date was adopted from Sacks et al. (2010) and Kucharik (2008), and the temperature threshold is chosen to ensure that soil temperature will be heated over 10^{o}C, which is normally the lower temperature threshold for crop growth. As soybean is traditionally sowed 15 days after maize, the new planting date for soybean is chosen as 15 days after the new planting date for maize for each year, GCM and RCP.

To estimate the impacts of climate change on crop growing cycles, the four growth stages in the future is estimated based on the average historical GDDs, which in turn is interpolated based on the FAO-tabulated growth stage lengths (in calendar days) and the typical growing season GDDs for maize and soybean according to (see Table 4-1). Every year, from the planting date, each crop growth stage lasts from the first day to the day when the growth stage total GDD is reached. Total growth stage GDD is the accumulated daily GDD, which is estimated in the following equation

$$GDD_j = \max(T_{mean} - T_{lower}, 0) \text{ when } T_{upper} > T_{mean}$$
$$\text{or} \quad \textit{Eq. 4-3}$$
$$GDD_j = \min(T_{upper} - T_{lower}, T_{mean} - T_{lower}) \text{ when } T_{upper} < T_{mean}$$

$$T_{mean} = \frac{T_{max} - T_{min}}{2} \quad \textit{Eq. 4-4}$$

Where: T_{max}, T_{min} and T_{mean} are maximum, minimum and average daily temperature during the growing season, T_{lower} is the lower threshold for crop growth temperature, below which GDD is equal to 0, and T_{upper} is the upper threshold, above which no GDD contributes to growth (in this study, 10 and 30°C were selected for the lower and upper thresholds, respectively, for both maize and soybean).

4.2.1.1.3 Effects of elevated CO_2 concentration

Studies have shown that elevated atmospheric CO_2 concentration ($[CO_2]$) effects crop water needs through reduction in stomatal conductance and transpiration per unit of leaf area. Decline in stomatal conductance in various plants due to elevated $[CO_2]$ results in a reduction of transpiration (Morison 1985, Allen Jr. 1990, Idso et al. 1994, Norby and Zak 2011). Experiments have been conducted to estimate the rate of change in stomatal conductance, transpiration efficiency, etc. in crops to varied $[CO_2]$. For example, studies by (Morison and Gifford 1983, Morison 1987) showed that when the $[CO_2]$ doubles, conductance would be linearly reduced to about 60% of current values between 330 ppm and 660 ppm; stomatal conductance would reduce by 41%, 47% and 39% for soybean, sweet corn and sweetgum, respectively, at $[CO_2]$ of 660 ppm from 330 ppm (Rogers et al., 1983). Therefore, there have been various methods to adjust crop ET to elevated levels of CO2 is to vary the stomatal conductance, which leads to changes in the bulk surface resistance (r_s) via the Eq 4-5:

$$r_s = \frac{r_l}{LAI_{active}} \qquad \textit{Eq. 4-5}$$

Where

- LAI_{active} is active (sunlit) leaf area index (m^2 (leaf) m^{-2} (soil)). For the reference crop, LAI_{active} is estimated as:

$$LAI_{active} = 0.5 * LAI_{CO2} \qquad \textit{Eq. 4-6}$$

LAI is adjusted to the elevated CO2 levels by the following formula:

$$LAI_{CO_2} = LAI[1 + \frac{7}{100}(\frac{[CO_2] - 330}{330})] \qquad \textit{Eq. 4-7}$$

Where and $[CO_2]$ is the concentration of CO_2. For reference grass, LAI = 24*h, where h is the height of the reference grass (0.12 m).

- r_l is bulk stomatal resistance of the well-illuminated leaf (s m^{-1}):

$$r_l = \frac{1}{g_{co_2}}$$ *Eq. 4-8*

There have been various equations to correct the stomatal conductance according to the [CO_2]. Stockle (1992) presented a linear equation to adjust g_{CO_2} to [CO_2] based on based on experimental observations of a 40% decrease in stomatal conductance between 330 and 660 ppm [CO2]. The approach of Morison and Gifford (1983) and has been used in several climate change impact studies (Easterling, Rosenberg et al. 1992, Parajuli 2010, Wu, Liu et al. 2011, Valenzuela, Johnson et al. 2016). However, Allen Jr (1990) pointed out that the elevation of [CO_2] by the end of the century may exceed the upper threshold in the aforementioned experiment, and that the stomatal conductance relationship with [CO_2] may be non-linear, similar to the biomass growth and [CO_2] relationship (Scarpare et al., 2022). Thus, in this study, the formula developed by Allen Jr (1990) was used (Eq. 4-9), based on an experiment in which [CO_2] was varied over a range from 340 to 1000 ppm.

$$g_{CO_2} = 0.0485 - 7.00 * 10^{-5}[CO_2] + 3.40 * 10^{-8}[CO_2]^2$$ *Eq. 4-9*

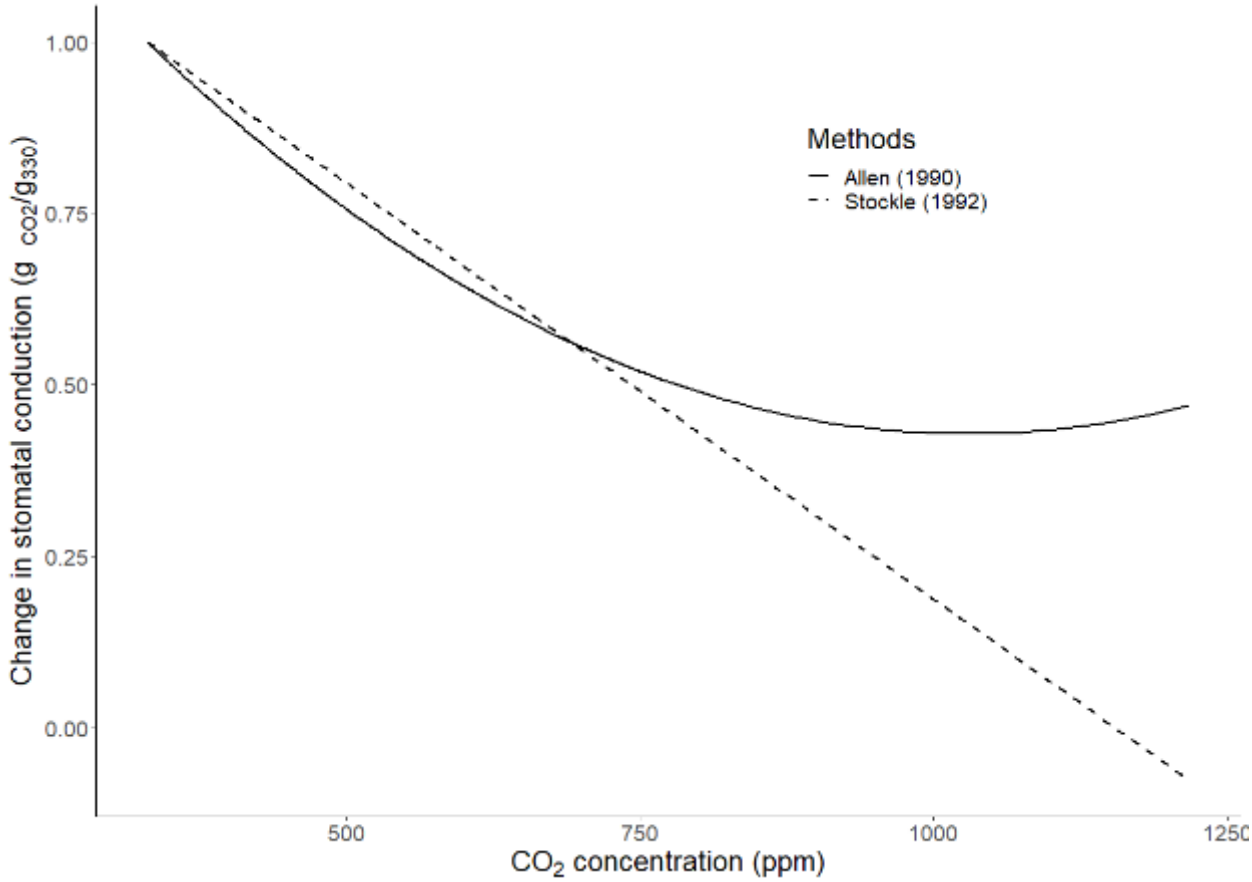

Figure 4-1 Changes in stomatal conductance according to formula by Stockle (1992) and Allen Jr (1990).

4.2.1.2 Effective precipitation

Effective precipitation Pe is the part of rainfall available to meet consumptive needs of crops, average monthly values. Scientists analyzed 50 years of rainfall records at 22 locations throughout the US to develop an equation to estimate EP (Martin and Gilley, 1993):

$$EP = 25.4 * SF * (0.04931 * P_t^{0.82416} - 0.11565) * (10^{0.000955 ET_c}) \quad \textit{Eq. 4-10}$$

where SF is the soil water storage factor:

$$SF = 0.531747 + 0.295167*D - 0.05797*D^2 + 0.003804*D^3 \quad \textit{Eq. 4-11}$$

where D is the usable soil water storage, estimated as 40-60% of soil available storage water (AWS) in the root zone depth. In this study, D is estimated at 50% of AWS at the 100 cm (soybean) and 150 cm depth (maize); ET_c is the crop potential evapotranspiration for a specific crop (see section 4.2.1).

4.2.2 Data

4.2.2.1 Climate data

An ensemble of 18 GCMs were taken from the University of Idaho Gridded Surface Meteorological Dataset under two climate change scenarios: RCP 4.5 and RCP 8.5 were used to estimate ET_c and PE. Changes in the climate under these two scenarios are presented as the region-mean, ensemble-mean, with standard deviation in Section 4.3.1. The GCMs were bias-corrected and statistically downscaled with the MACA (Multivariate Adaptive Constructed Analogs) method to a spatial resolution 1/24th degree gridcell resolution (Abatzoglou, 2013). To avoid climate model biases, the results in this study were expressed as relative changes from the near future (2030–2059) and far future (2070-2099) to the baseline period (1980–2009).

Two [CO_2] scenarios were used in this study to examine the CFE on crop net irrigation need (see Figure 4-2): recent historical [CO_2] (referred to as *contemporary* [CO_2]) where [CO_2] is fixed at 387 ppm, and the elevated [CO_2] trajectories for two the RCPs by the IPCC (referred to as *elevated* [CO_2]) (Stocker, 2014).

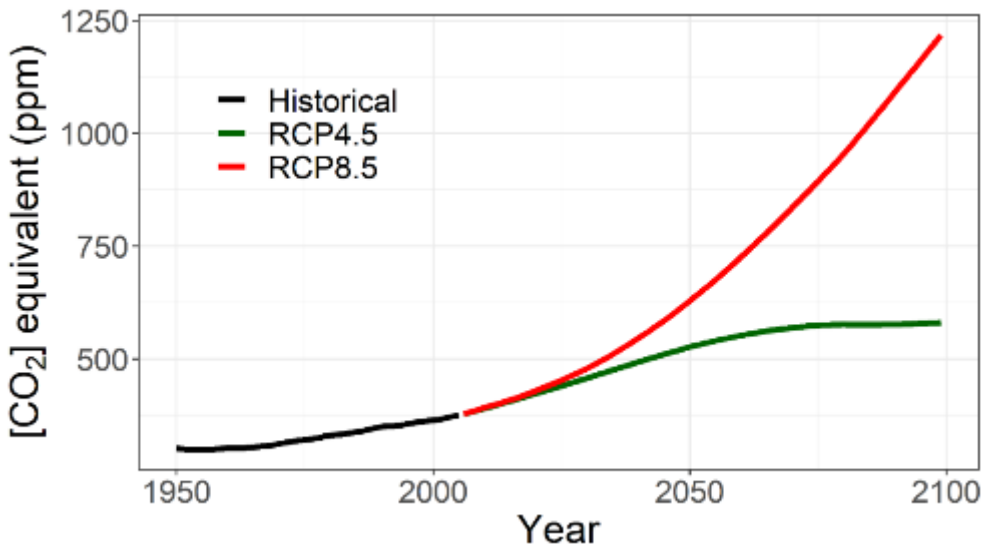

Figure 4-2 The [CO2] trajectories under RCP 4.5 and RCP 8.5 according to the IPCC

4.2.2.2 Soil data

Available soil water storage (AWS) is aggregated from the Gridded Soil Survey Geographic (gSSURGO) database[9]. The gSSURGO datasets are available at 10x10 m and 30x30 m resolution. The data were collected and aggregated to 1/24th degree resolution to match the spatial resolution of the meteorological inputs. The collection and aggregation of AWS were conducted via the R packages "soilDB" and "raster." AWS is available at fourth depths: 25, 50, 100 and 150 cm (see Figure 4-2). For maize, AWS at soil depth of 150 cm was used; for soybean, AWS at soil depth of 100 cm was used. Figure 4-2 shows that the western area of the region, where the Mississippi river is located, tend to have higher AWS than other areas.

[9] https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/home/?cid=nrcs142p2_053628

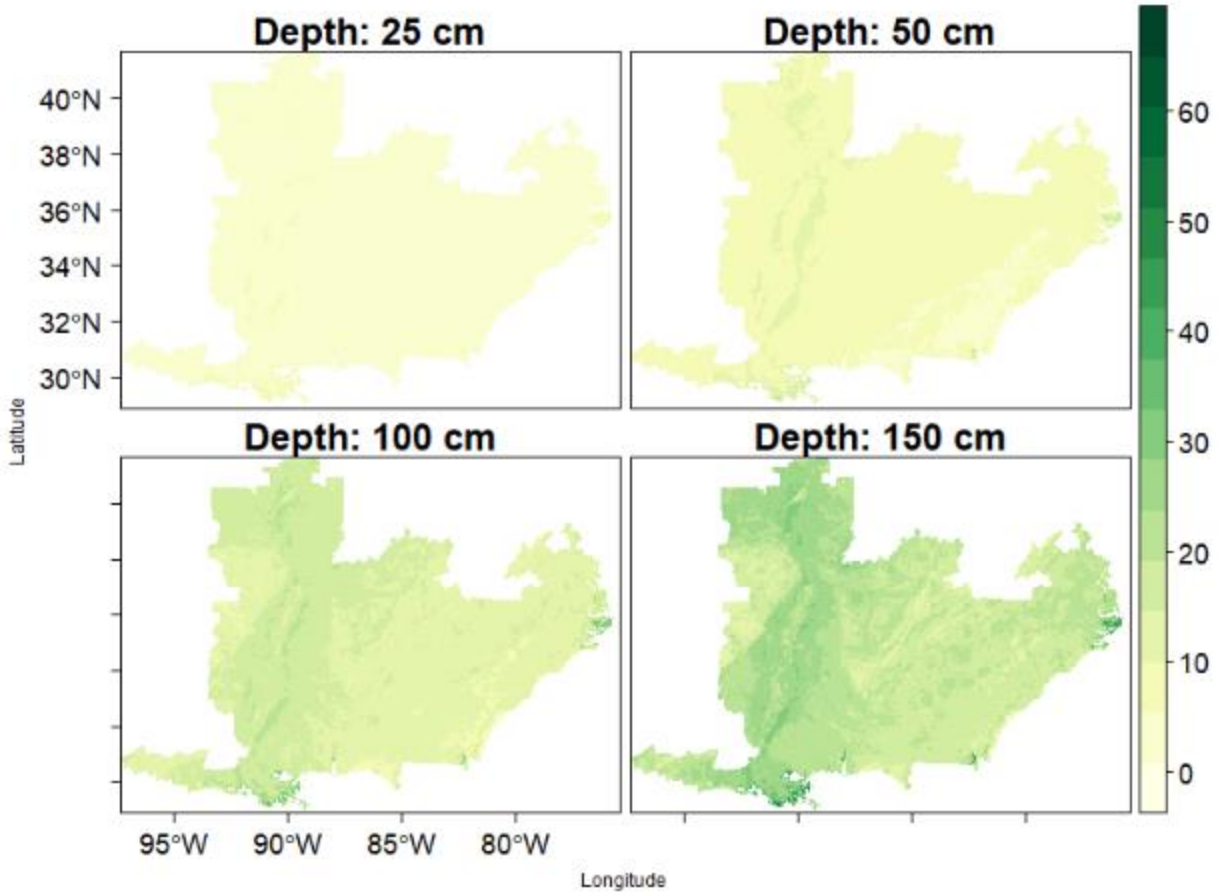

Figure 4-3 Available soil water storage (AWS) maps across the study region at four soil depths, aggregated from the Gridded Soil Survey Geographic (gSSURGO) database

4.3 *Results and discussion*

4.3.1 Changes in future climate

The results from an ensemble of GCMs show changes in precipitation (Pr) do not show a consistent pattern across the ensemble of GCMs compared to temperature, as show in Figures 4-4 and 4-5. Summer months may show a declining trend, especially under RCP 8.5, but the trends are not significant. Decline in effective precipitation (EP) tends to be smaller than Pr as lower rainfall may lead to lower runoff, thus lower water loss from the soil. In the month of July, the month that was identified as highly critical to the crops (see Chapter 3), Pr may decrease by 3.92 or 0.84 mm/day for the near future period under RCP 4.5 and RCP 8.5 respectively; for the far future period, the projections of Pr decrease are 1.31 or 8.71 mm/day under RCP 4.5 and RCP 8.5, respectively. EP in July is estimated to decline by 2.89 and 1.05 mm/day under RCP 4.5 and 8.5,

respectively, for the near future; the projections for the far future is 0.95 and 6.54 mm/day under RCP 4.5 and 8.5, respectively.

Maximum and minimum temperature (T_{min} and T_{max}) will very likely increase in the near (2030-2059) and far (2070-2099) futures, under both RCP 4.5 and RCP 8.5 (see Figures 4-6 to 4-7). The increases are highest during the summer time (April to September) which coincides with growing season for maize and soybean. In particular for July, the average increase in T_{max} over the study region is about 1.98ºC or 2.5ºC in the near future, and 2.69ºC or 5.32ºC in the far future, under the RCP 4.5 or RCP 8.5 projections respectively. The corresponding increases in T_{min} for July is 1.76ºC or 2.3ºC and 2.51ºC or 4.95ºC, again under the RCP 4.5 or RCP 8.5 respectively.

The increase in maximum and minimum temperature tends to have multiple effects on plant water uses, primarily through increasing ET_o. However, as shown in Figures 4-4 to 4-9, increases in T_{max} and T_{min} do not directly affect in ET_o, but through changes in vapor pressure deficit (VPD). At a given temperature, VPD is the difference between the saturation vapor pressure and the actual vapor pressure. Because saturation vapor pressure increases exponentially with temperature, VPD increases exponentially when the difference between maximum and minimum temperature increases and absolute humidity remains constant. Figures 4-8 and 4-9 show that the seasonal pattern of ET_o follow the pattern of VPD. VPD and ET_o will increase across all months but the highest increases occur in the summer time. Under the RCP 4.5 and RCP 8.5, respectively, July ET_o will increase by 8.30 and 12.37 mm/day for the near future period, and 8.48 and 23.64 mm/day for the far future period.

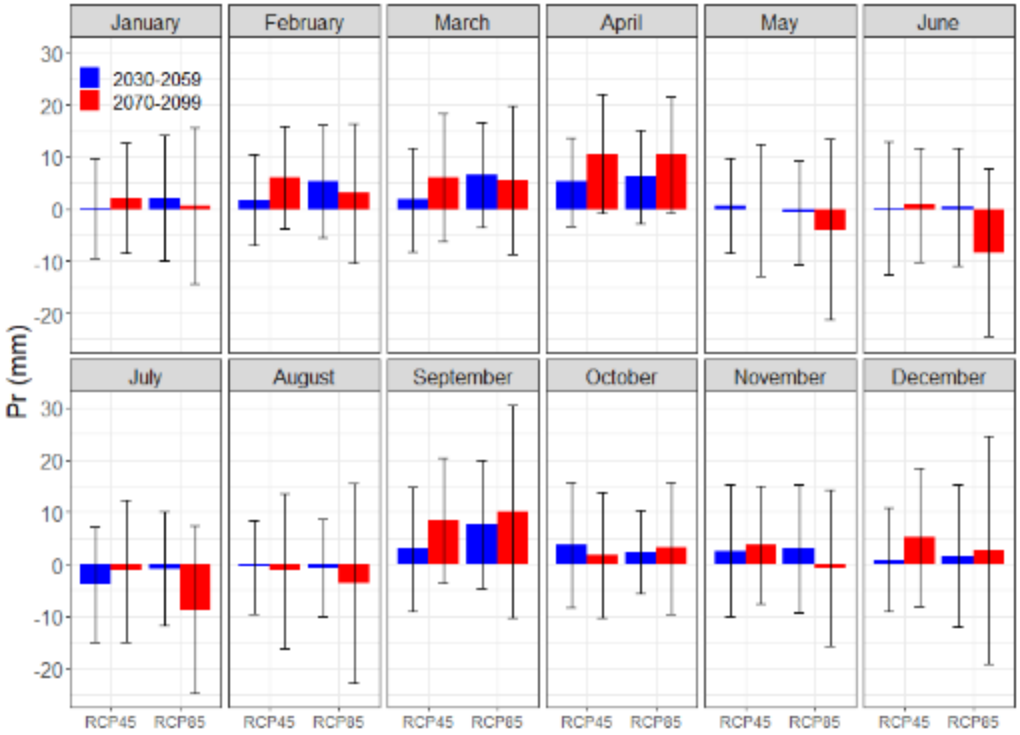

Figure 4-4 Changes in precipitation (Pr) under future climate change scenarios RCP 4.5 and RCP 8.5 for the near and far future periods compared to the baseline (1980-2009). The columns represent the mean change across the region and the ensemble of GCM, and the whiskers represent the standard deviation of the mean change.

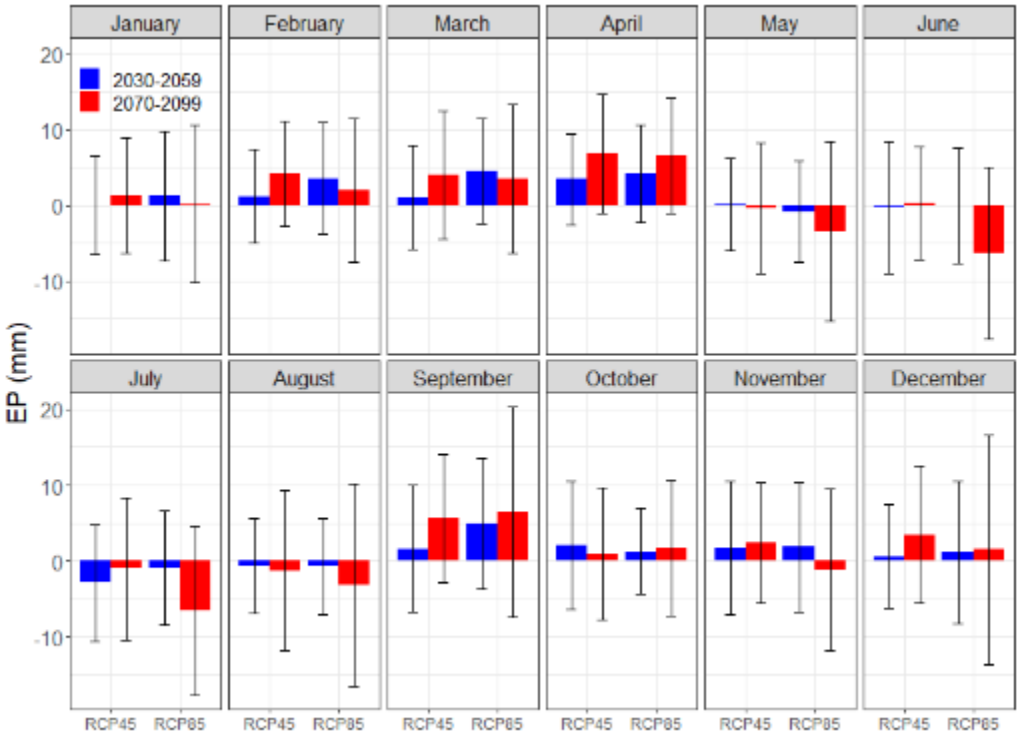

Figure 4-5 Similar to 4-4, but for effective precipitation (EP)

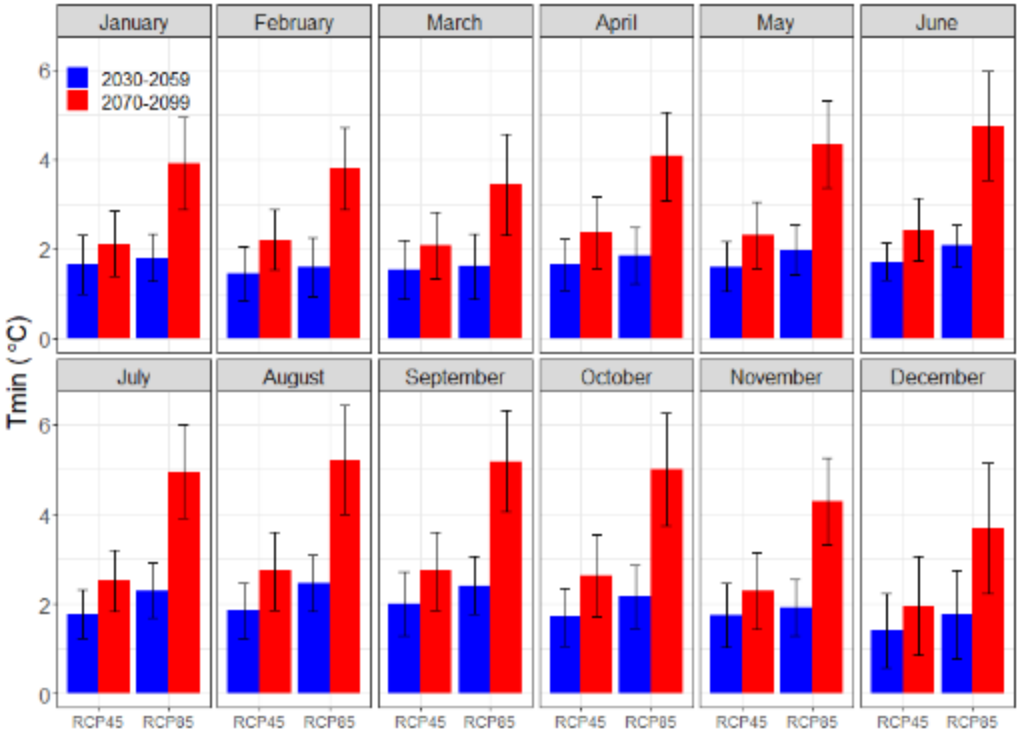

Figure 4-6 Similar to 4-4, but for minimum temperature (T_{min})

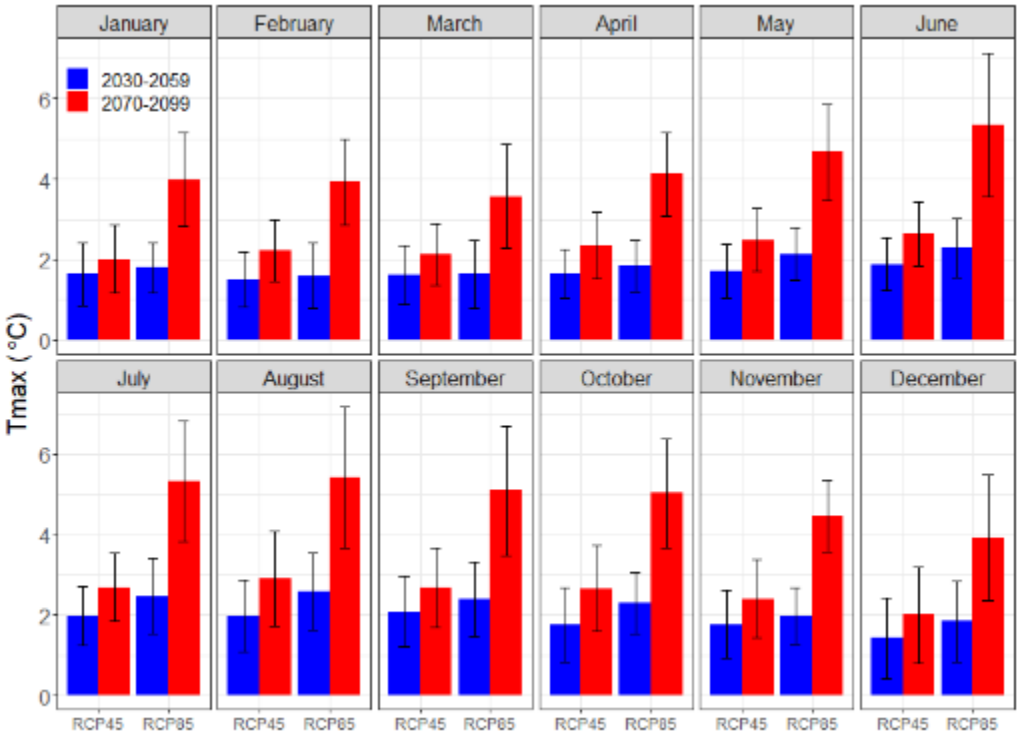

Figure 4-7 Similar to 4-4, but for maximum temperature (T_{max})

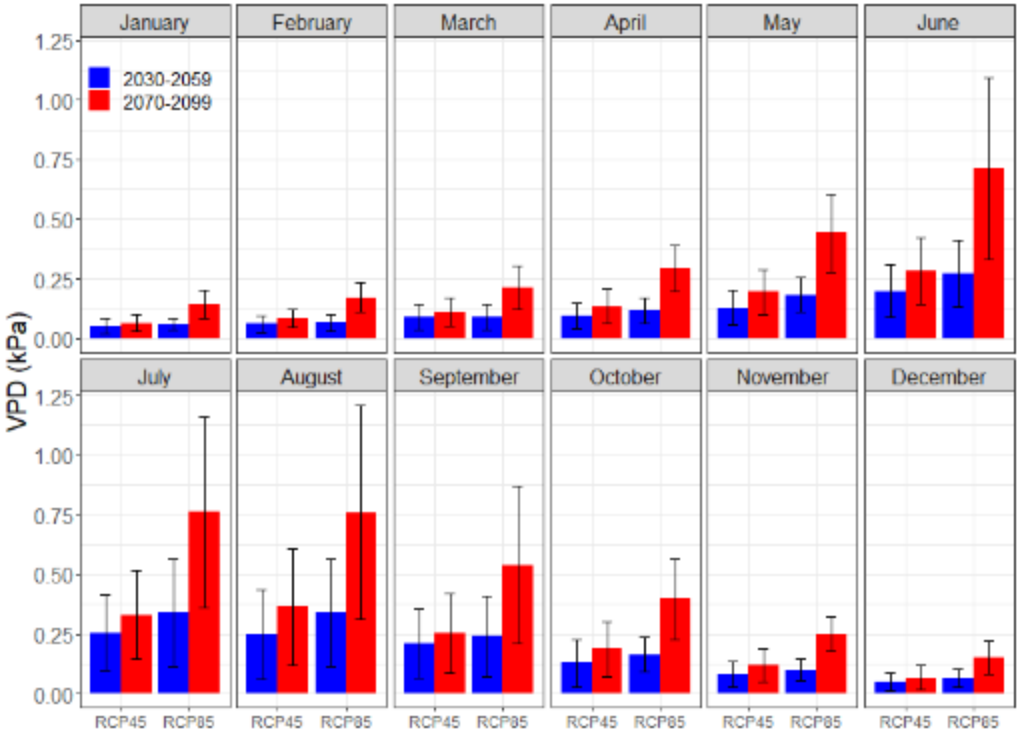

Figure 4-8 Similar to 4-4, but for vapor pressure deficit (VPD)

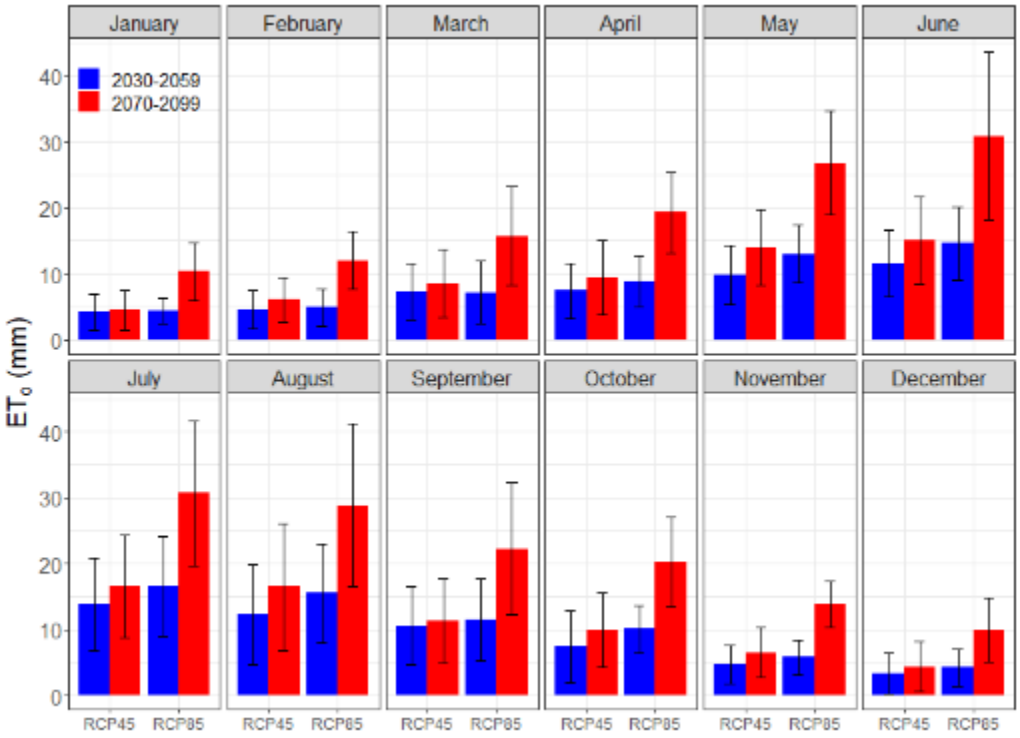

Figure 4-9 Similar to 4-4, but for reference evapotranspiration (ET_o)

4.3.2 *Changes in crop potential evapotranspiration (ET_c)*

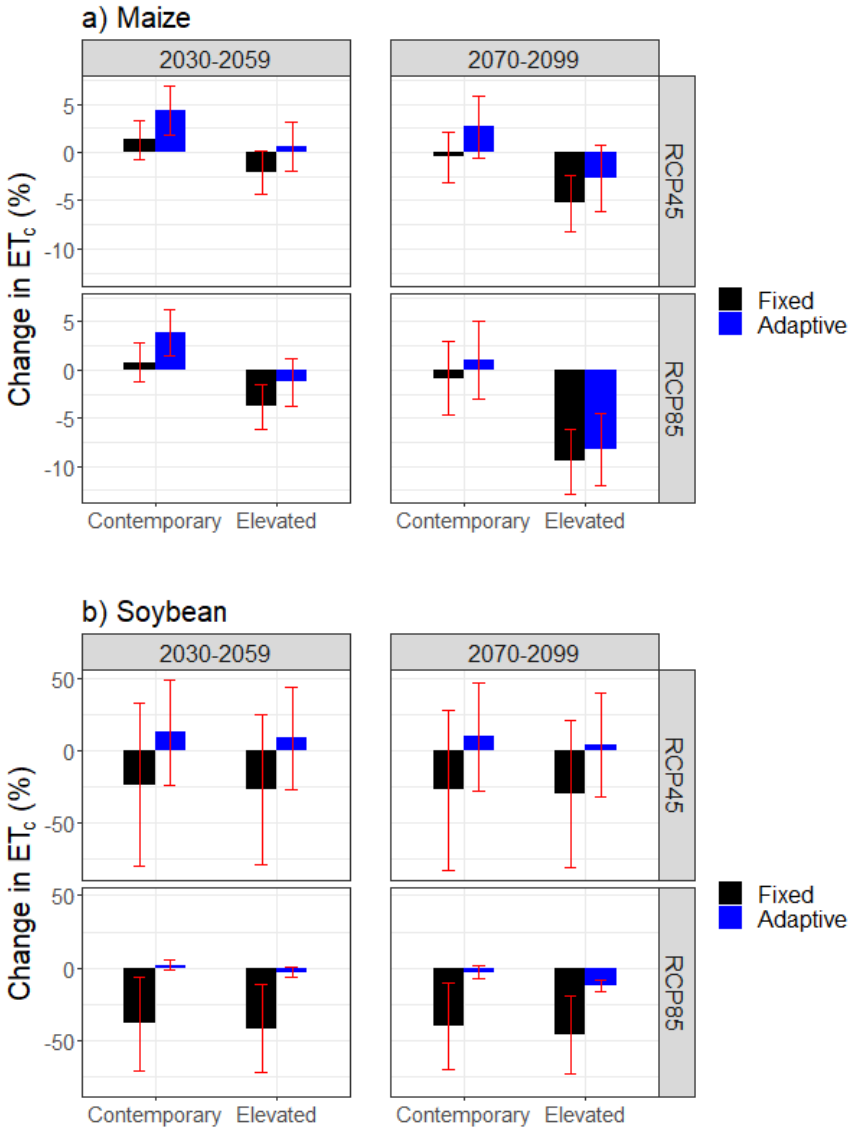

Figure 4-10 Changes in crop potential evapotranspiration (ET_c) under future climate change scenarios RCP 4.5 and RCP 8.5 for the near and far future periods compared to the baseline (1980-2009) for two planting date scenarios: fixed planting date (black columns) and adaptive planting date (blue columns). The horizontal axis represents results from two [CO_2] scenarios: Contemporary where future [CO_2] is fixed at 387 ppm, and Elevated to represent changes in [CO_2] in the future. The columns represent the mean change across the region and the ensemble of GCM, and the whiskers represent the standard deviation of the mean change.

Changes in ET_c (expressed as % of changes compared to the baseline period) (Figure 4-10) for maize and soybean under future climate differ according to the planting date scenarios and whether the CFE is included. Elevated [CO_2] generally has a positive effect on crop water demand: with the same planting date, when the CFE was considered, ET_c was lower under all RCPs and for both crops. In general, under the fixed planting date, crop water demand tends to be lower

compared to the adaptive planting date. For maize, under the fixed planting date, the growing season length tends to be 3 to 5 days shorter in the near future, and 7 to 15 days shorter in the far future under RCP 4.5 and RCP 8.5 scenarios for maize. For soybean, the corresponding average growing season length is both reduced by 7 to 12 days and 15 to 20 days. If farmers start planting at the historical date, there is a risk that growing season will be shortened due to accelerated maturity as a result of fastened accumulation of growing season GDD. This will lead to an overall reduction in biomass accumulation and yield. Generally, a shorter growing season leads to lower ET_c, except for maize in the near future under RCP 4.5, as the slight shorter growing season does not compensate for increase climatic evaporative demand from increased temperature.

The adaptive planting date takes advantage of warmer spring time, and reduces the risk of accelerated maturity. On average, the growing season may start 5 to 10 days earlier in the adaptive planting date scenario compared to the historical date in both the near and far future for maize and soybean under both RCPs. Under RCP 8.5, the growing season length with an adaptive planting date is on average 7 days shorter than the historical length for both maize and soybean, while under the fixed planting date scenario, the growing season length can be reduced by 19 and 25 days on average for maize and soybean, respectively. Under RCP 4.5, the growing season length with an adaptive planting date may increase by 6 and 15 days on average for maize and soybean, while approximately 10 days shorter with a fixed planting date for both crops. This increase in the growing season length under the moderate warming scenario leads to increase in crop water demand for adaptive planting date scenario compared to the fixed planting date.

The risk of increased temperature is much higher for soybean than maize. This is likely attributed to the fact that soybean require smaller GDDs to reach maturity than maize, thus the increase in temperature tends to fasten soybean to maturity faster than maize, thus the growing season of soybean tends to be shorter than maize in the future. Shorter growing season due to increase temperature will reduce soybean biomass accumulation and yield, make it more vulnerable to climate change than maize. Therefore, adaptive planting date will help to reduce the risk of heat exposure to maize and soybean production.

4.3.3 Changes in net irrigation requirement

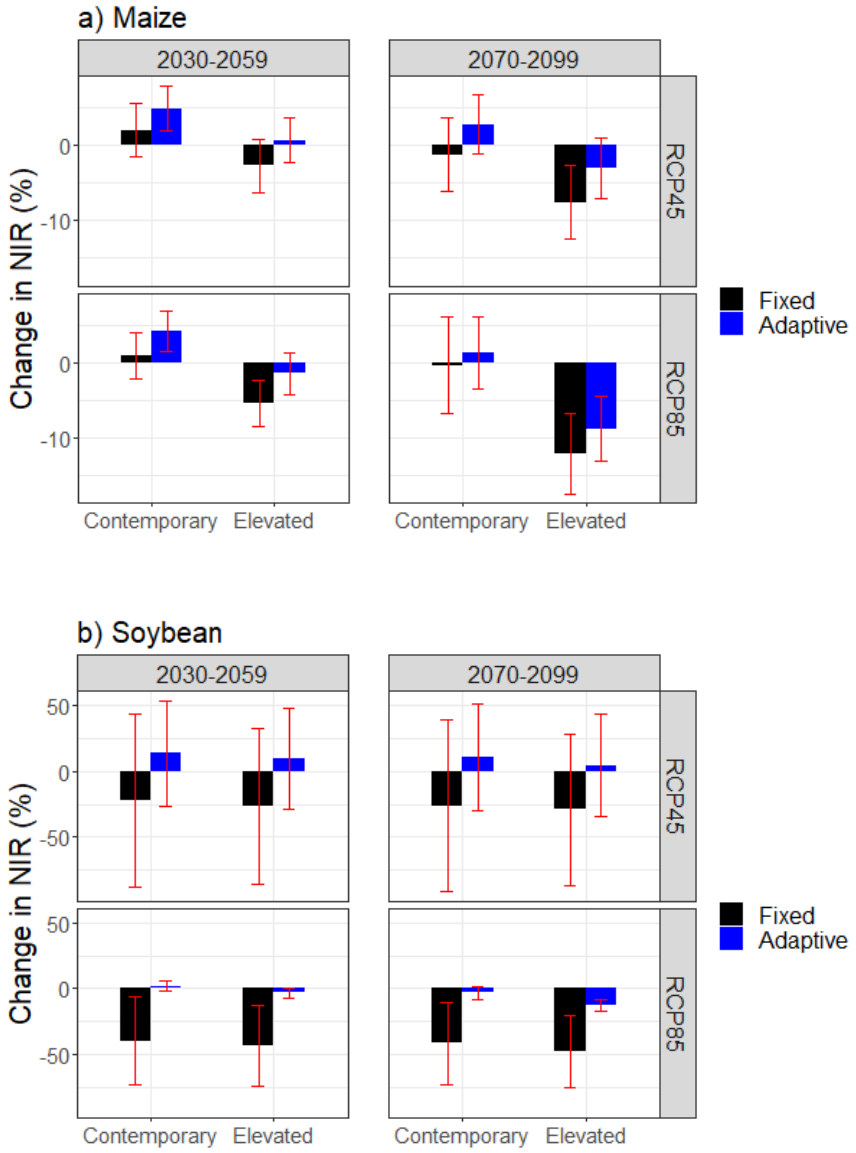

Figure 4-11 Changes in net irrigation requirement (NIR) under future climate change scenarios RCP 4.5 and RCP 8.5 for the near and far future periods compared to the baseline (1980-2009) for two planting date scenarios: fixed planting date and adaptive planting date. The horizontal axis represents results from two [CO_2] scenarios: Contemporary where future [CO_2] is fixed at 387 ppm, and Elevated to represent changes in [CO_2] in the future. The columns represent the mean change across the region and the ensemble of GCM, and the whiskers represent the standard deviation of the mean change.

Changes in net irrigation requirement (NIR) is the relative change in the difference between the supply of water (EP) and the crop water demand (ET_c). Figure 4-11 shows the ensemble-mean change of NIR across the study region for maize and soybean. A positive change in NIR indicates that the net water requirement is going to increase, while a negative value indicates a reduction in the requirement. In general, NIR follows a similar pattern to ET_c, since changes in EP is rather small and do not have a clear pattern to affect the pattern of NIR.

The fixed planting date tends to reduce NIR compared to the historical baseline due to a shortened growing season; the exception is NIR of maize in the near future under RCP 4.5 and the contemporary [CO_2] scenario, when the impact of slightly shortened growing season on NIR is offset by higher increase in crop water demand due to warmer climate. In the near future, without the CFE, NIR of maize will change by 1.94 and -2.75 % under RCP 4.5 and RCP 8.5, respectively; with CFE, the changes in NIR are projected to be -1.31 and -7.65 % under RCP 4.5 and RCP 8.5, respectively. For the far future, NIR of maize will change by -3.42 and -1.91 % under RCP 4.5 and RCP 8.5, respectively under contemporary [CO_2]; under elevated [CO_2], NIR is projected to decline by -12.2 and -15.6 % under RCP 4.5 and RCP 8.5, respectively. For soybean, NIR changes under contemporary [CO_2] scenario are projected to be -22.1 and -39.6 % in the near future, and -25.2 and -41.9 % in the far future, under RCP 4.5 and RCP 8.5, respectively; NIR changes under the elevated [CO2] scenario are estimated at -29.29 and -43.22 % in the near future, -30.44 to -47.57 % under RCP 4.5 and RCP 8.5, respectively.

With adaptive planting date, the risk of accelerated maturity is reduced, but NIR tends to increase under this scenario as the climatic demand for evapotranspiration increases due to higher temperature. However, even with CFE taken into consideration, high level of warming such as under RCP will have strong negative impacts of crop production, as extreme heat will shorten growing season and reduce yield. In the near future, NIR of maize will change by +5.82 and +3.37 % under RCP 4.5 and RCP 8.5, respectively, at contemporary [CO_2]; these changes are projected to be +2.76 and +0.64 % under RCP 4.5 and RCP 8.5, respectively at elevated [CO_2]. For the far future, NIR of maize will change by +4.18 and +1.37 % under RCP 4.5 and RCP 8.5, respectively under contemporary [CO_2]; under elevated [CO_2], NIR is projected to change by -3.09 and -8.85 % under RCP 4.5 and RCP 8.5, respectively. For soybean, NIR changes under contemporary [CO_2] scenario are projected to be 13.53 and 4.51 % in the near future, and 9.10 and 2.05 % in the far future, under RCP 4.5 and RCP 8.5, respectively; NIR changes under the elevated [CO_2] scenario are estimated at 10.57 and 3.23 % in the near future, 3.11 to -12.36 % in the far future, under RCP 4.5 and RCP 8.5, respectively.

Figures 4-12 to 4-15 show that the spatial distribution of NIR will likely be in the northern (especially northeastern) area of the study region, whereas the southern areas of the study region will experience little change or decrease in NIR. Here results from RCP 4.5 are shown for brevity,

and because this scenario shows higher potential of increase in NIR. The southern part of the study region is historically warmer, thus increased temperature will lead to fastened accumulation of GDD, especially under fixed planting date. Therefore, this region will experience shorter growing season and no change or lower NIR compared to the northern region, where cooler historical temperature allows slower accumulation of GDD. Most of the crops are grown in the northwestern, southwestern and southeastern area of the study region. As results from Chapter 2 shows, there is significant increase in mean temperature in the future for these areas, which subsequently alter the crop growing cycle and NIR. Therefore, increase in NIR in these regions need to be taken into consideration for future planning. Under a fixed planting date scenario, these crop-growing areas may experience between -20.1 % and +20.2 % change in NIR for maize, and a 0 to +60.1 % increase for soybean. Under adaptive planting date, NIR may increase much more as the crops may maintain its growing season length, thus requires higher ET_c, the crop-growing areas may experience up to 80% increase in NIR for maize and 100% increase for soybean, especially in the northeastern area. As this region is traditionally rainfed, investment on irrigation infrastructure may be necessary to insure crop productivity in the future.

Maize - Fixed planting date

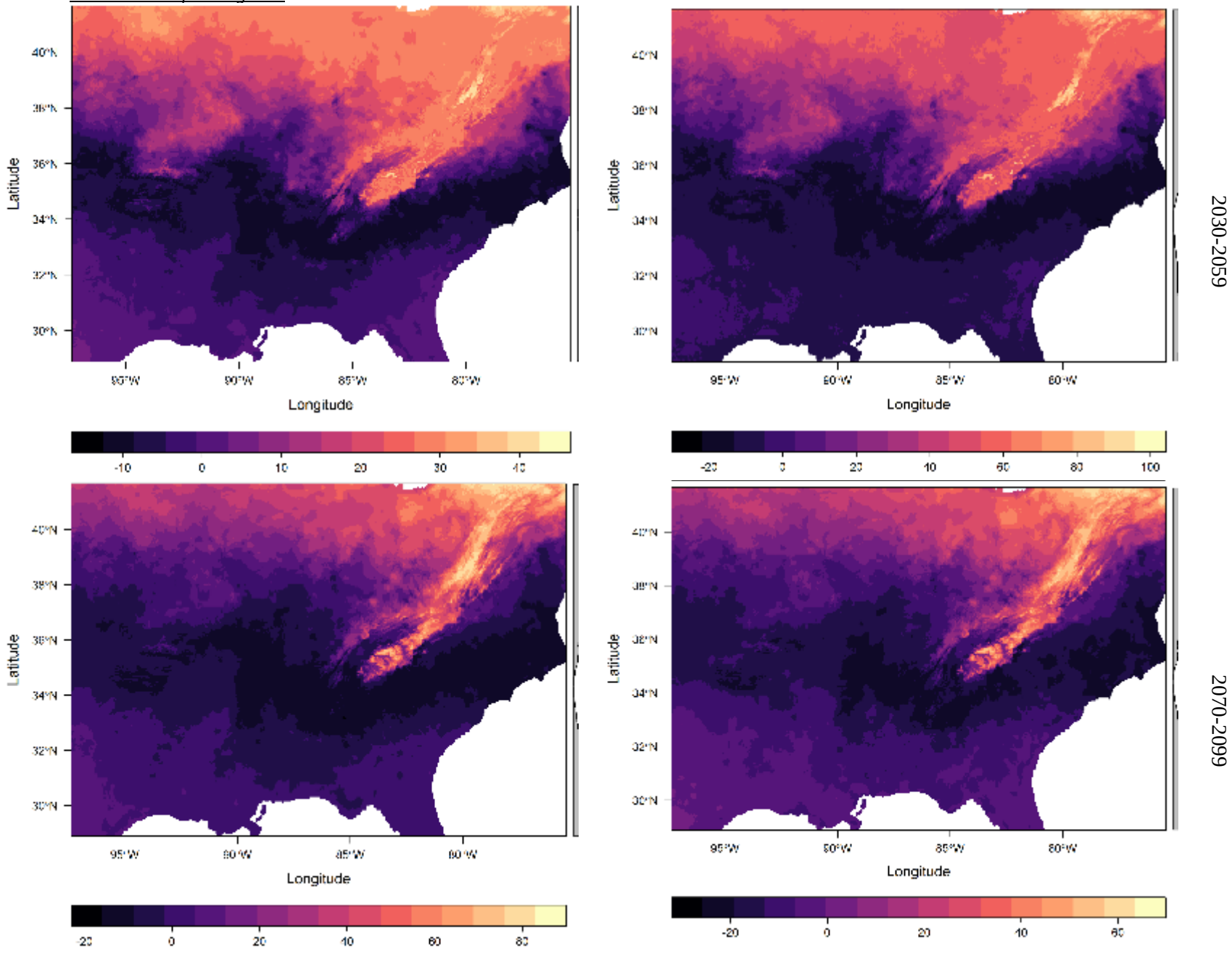

Figure 4-12 Spatial pattern of net irrigation requirement (NIR) for maize with a fixed planting date under RCP 4.5. NIR is expressed as the relative change (%) from the future periods to the baseline period (1980 to 2009)

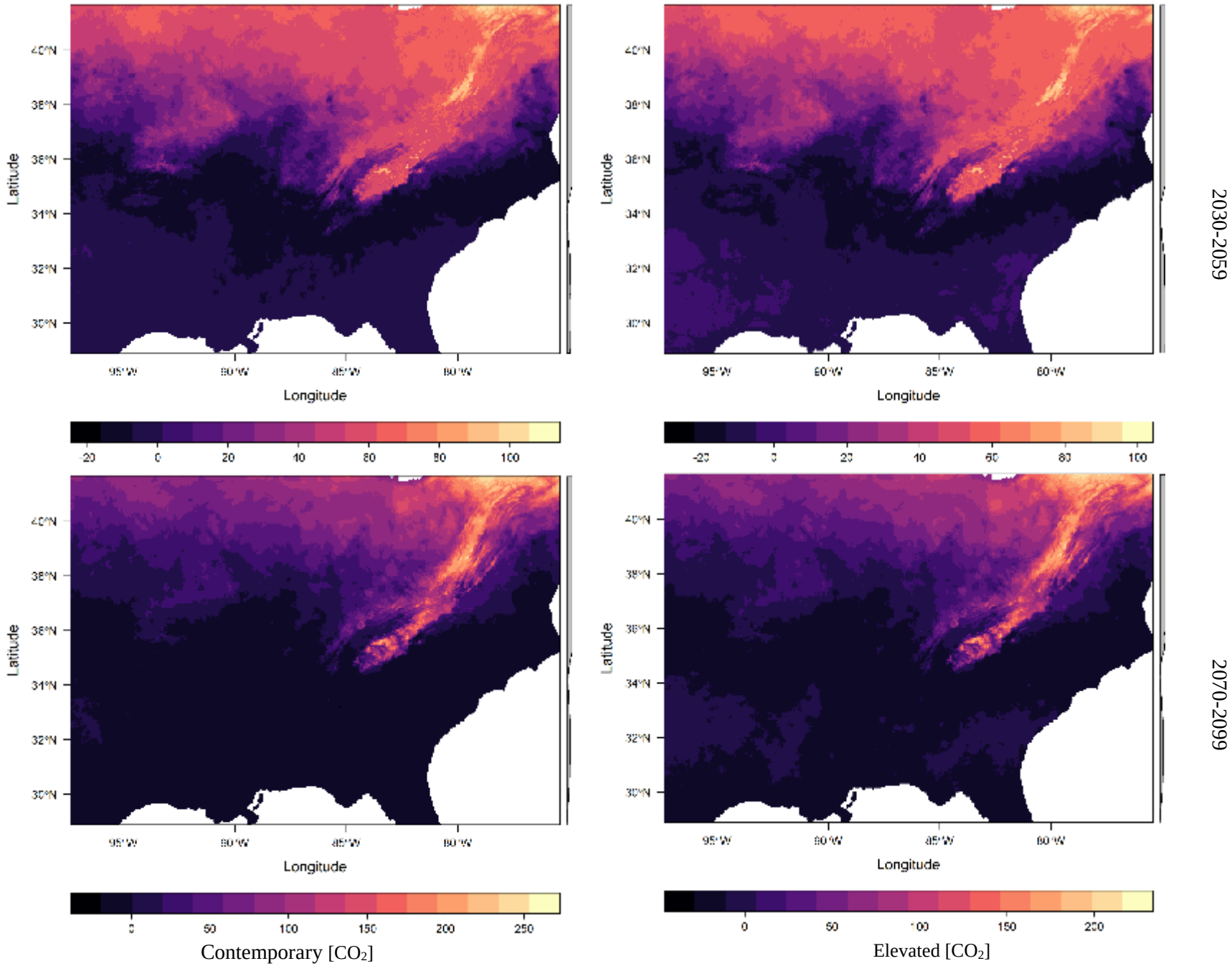

Figure 4-13 Similar to 4-12 but for maize with an adaptive planting date

Soybean – Fixed planting date

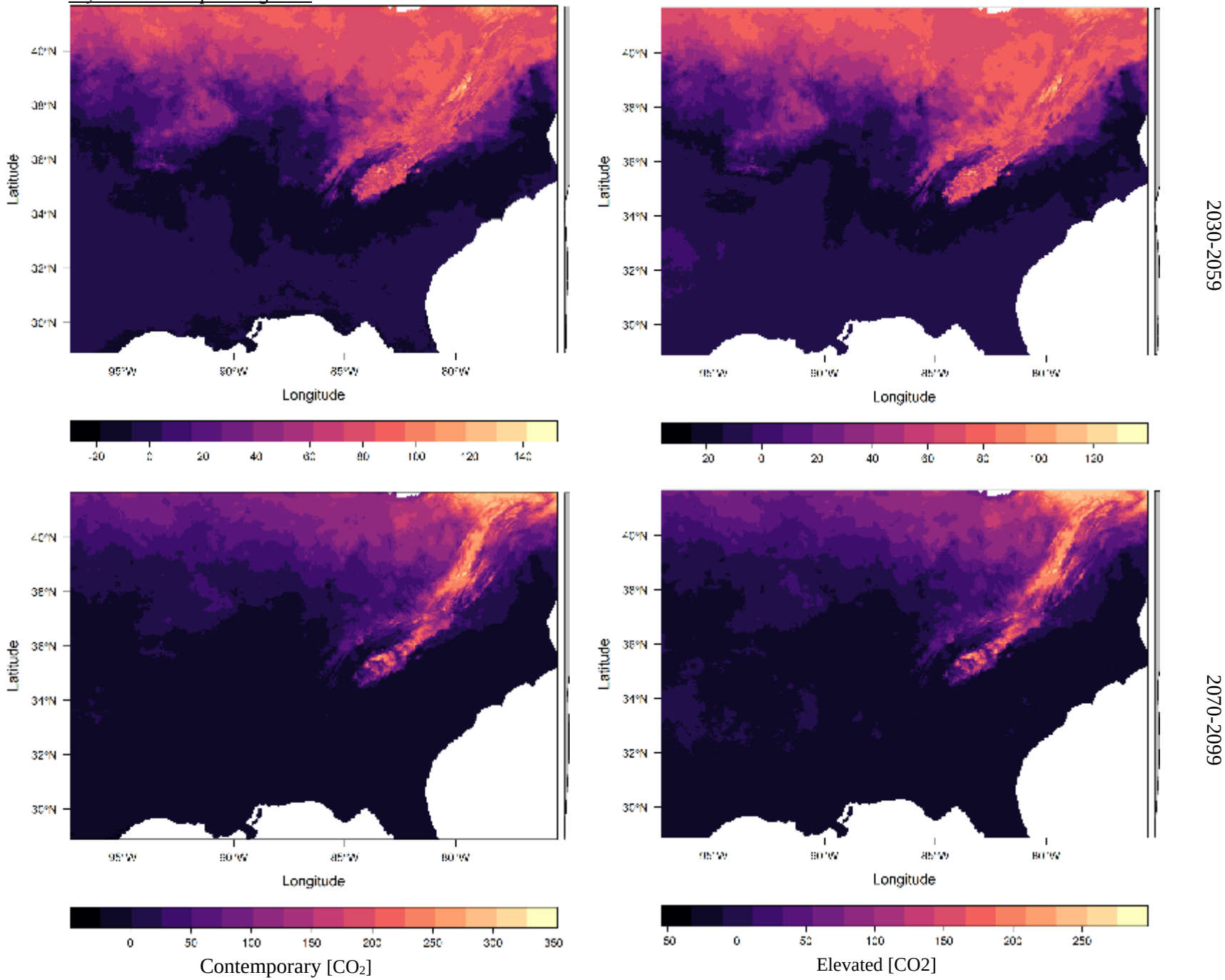

Figure 4-14 Similar to 4-12 but for soybean with a fixed planting date

Soybean – Adaptive planting date

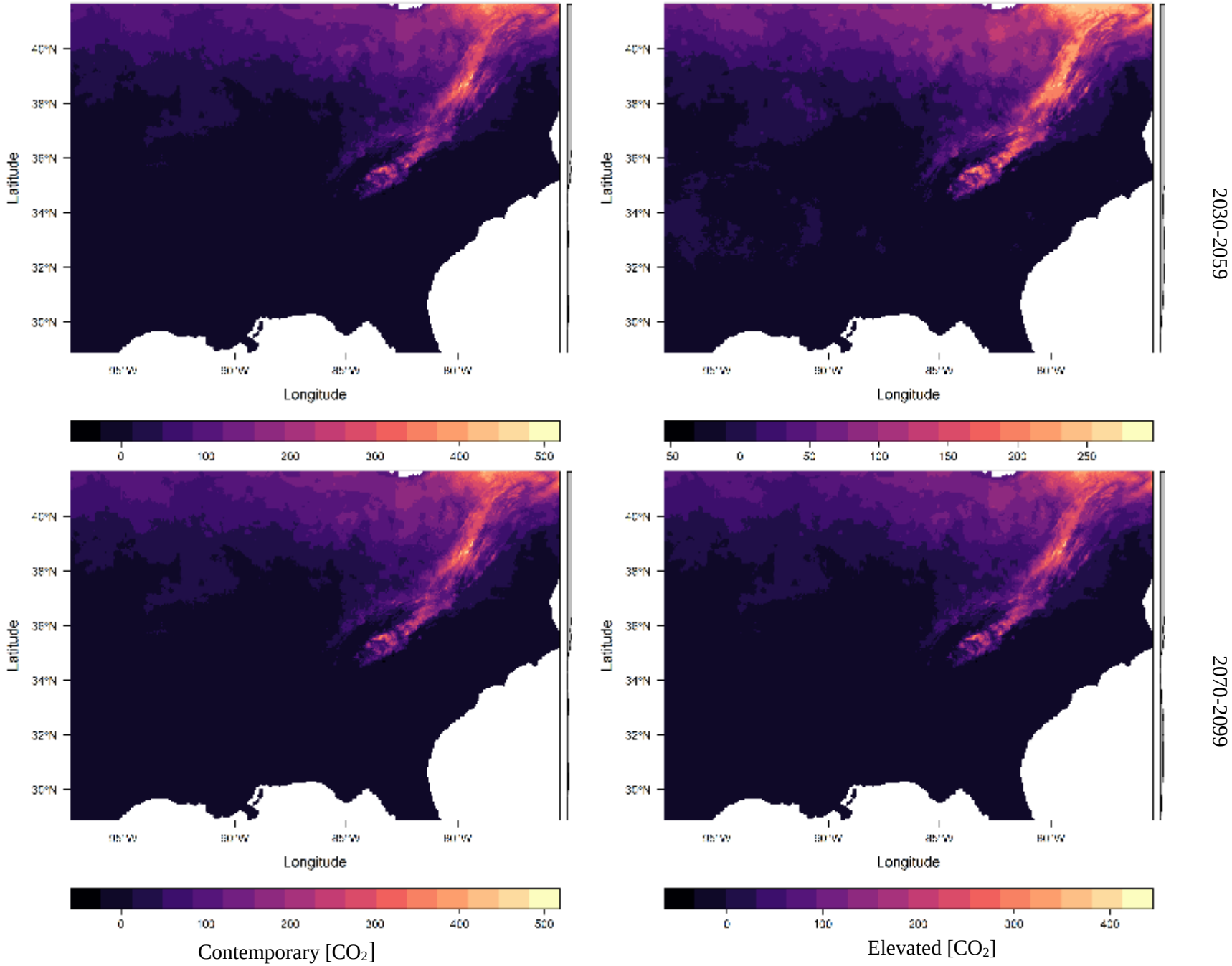

Figure 4-15 Similar to 4-12 but for soybean with an adaptive planting date

4.4 *Conclusions*

Crop potential evapotranspiration and net irrigation requirement estimations from an ensemble of 18 downscaled GCM that show that there can be a wide range of projected changes in ET_c and NIR as compared to the historical baseline (1980 to 2009). ET_c and NIR estimations follow similar seasonal pattern mostly because changes in effective precipitation in the region are not large enough to offset the changes in ET_c. These changes depend on crop type, CFE assumptions and planting date scenario. Crop growing cycle is altered by changes in temperature, and generally crop growing season will be shortened if crops are grown around the historical planting date. Under the fixed planting date, maize growing season length tends to be 3 - 15 days shorter in the future periods under RCP 4.5 and RCP 8.5; for soybean, the growing season length is potentially reduced by 7 - 20 days. With an adaptive planting date, the growing season may be 5-10 days shorter than the historical length under RCP 8.5, but may be 19-25 days longer the same under RCP 4.5 for maize and soybean, respectively. Fixed planting date increase the risk of heat exposure, fastened maturity, which will lead to an overall reduction in biomass accumulation and yield. Generally, shorter growing season leads to lower ET_c, except for maize in the near future under RCP45, as the slight shorter growing season does not compensate for increase climatic evaporative demand from increased temperature. With adaptive planting date, the risk of shortened growing season is reduced, thus the water requirement is enhanced under this scenario compared to a fixed date. Under the same planting date and RCP scenario, the effects of elevated CO_2 level may reduce crop NIR. The study finds that NIR tends to decrease under the fixed planting date scenario compared to the historical baseline due to a shortened growing season, and increase under the adaptive planting date scenario as the climatic demand for evapotranspiration increases due to higher temperature while the risk of accelerated maturity is reduced. The effects of elevated CO_2 level may reduce crop NIR under the same planting date scenario and RCP. Under the same planting date scenario and RCP, if CO_2 level is fixed at the contemporary level, NIR of maize will increase by 5.82 and 3.33 % in the near future and 4.18 and 1.37 % under RCP 4.5 and RCP 8.5, respectively, while NIR of soybean will increase by 13.53 and 4.51 % in the near future, and 9.10 and 2.05 % in the far future under RCP 4.5 and RCP 8.5, respectively. Spatial pattern of NIR shows that the southern area may experience little change or reduction in NIR due to increase in crop water demand under all planting date or NIR scenario as the impacts of increased temperature on shortened growing season length may overcome or cancel out enhanced climatic evaporative

demand. Of the crop-growing areas, the northwestern area near the Corn Belt, where crops were traditionally rainfed, may require investments in irrigation infrastructure, as this area may experience increase in NIR for both maize and soybean.

Chapter 5 Assessing climate change impacts on maize and soybean production through an ensemble of emulators and statistical models

Abstract

Understanding climate change impacts on crop production is crucial to anticipating future production and designing appropriate adaptation plans. One of the common approaches for analyzing yield response to climate is to develop empirical models using historical observations of yield and climate. This, however, has a limitation of extrapolating yield response to climate conditions outside the historical range (assumption of stationarity). Emulators, which use simulated climate and yield data from process-based models as 'true' observations to develop statistical models, are used to provide more insight about possible yield responses to climate change. In this paper, two different datasets were used to develop an ensemble of statistical models. The first set of data is historical crop yields and climate data, taken from the USDA-NASS database and gridMET datasets respectively, to construct a set of statistical models based on actual observations (called *observation model*). The second set of data are simulations taken from the Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP) to develop several *emulators*. These data are comprised of results from the three different global gridded crop models (GGCMs) with climate input from three bias-corrected CMIP5 global circulation models (GCMs) with the representative concentration pathway (RCP 8.5). This study focused on estimating of yield changes under future climate scenarios in the southeastern United States and aiming to analyzes differences in how the emulators simulating yield response to climate compared to the observation model. The emulators and observations show similar direction for the response of yield to changes in temperature and precipitation, but vary in the magnitude. The disparity in yield response to climate as a whole and to each climate component, i.e., precipitation and temperature, between the *observation models* and *emulators* are due to differences in the structure of the GGCMs as well as in the climate data. The results from these models show that under the impacts of climate change, median yield in the study area is expected to decline from 5% to 44% for maize and 3% to 78% for soybean in the near future. For the far future, the decline in yield is expected to be 25% to 72% for maize, and 11% to 96% for soybean.

5.1 *Introduction*

Despite technological advances, crop production is still susceptible to climate conditions (Gornall et al., 2010). A recent study found that there is an increasing risk of simultaneous crop failure as the global average temperature increase, which would put more pressure on the impoverished portion of the world population (Wheeler and von Braun, 2013). As future climate is projected to change (Stocker, 2014), understanding how these changes affect crop yield is important for planning and adaptation (Rosenzweig et al., 2014). Two common approaches have been used to study climate impacts on yield: i) process-based modeling, which simulates crop growth processes on a daily or even sub-daily basis as a result of detailed interactions between various physiological, agronomic and soil-water processes (Jones et al., 2003; Keating et al., 2003; Muchow et al., 1990; Stöckle et al., 2003; Tao et al., 2009), and ii) statistical modeling, which estimate econometric reduced-form responses of crop yield to climate factors (Auffhammer et al., 2011; Lobell et al., 2005; Lobell and Asner, 2003; Lobell and Burke, 2010; Nicholls, 1997; Schlenker and Lobell, 2010; Schlenker and Roberts, 2009). These two approaches both have their advantages and disadvantages that contribute to uncertainty in prediction of yield response to climate and must be considered when interpreting results of yield response to climate.

Process-based models are a set of mathematical equations developed based on laboratory experiments or field trials in a particular environment and time period. These equations represent the relationships observed between various crop growth factors such as climate, soil type, management practices, crop physiology, etc. Process-based models compute crop physiological response to climate conditions in an explicit, dynamic manner under a given management and technological practices and over a range of small to large spatial and temporal scales. In general, they can act as a digital laboratory to evaluate potential crop yield in response to a single or multiple factors (Lobell et al., 2009). While the level of sophistication and computation capacity of these models have increased, whether these models can represent the real-world growing conditions is still a challenge to modelers. Most process-based models do not account for factors such as pests, weeds or farmer behaviors to short-terms shocks that can affect yield formation (Lobell et al., 2009; Roberts et al., 2017). Assumptions about crop growth processes in process-based models are often limited to the data collected in the location used to calibrate the model. Therefore, large-scale application of these models outside the original location where it was developed is subject to uncertainty in terms of model parameterization and capability to capture

the response of crop yields (Roberts et al., 2017). In studying climate change impacts on yield using process-based models, various studies point out the need for using an ensemble of crop models because the ensemble mean or median tends to more closely approximate the historical value than any individual model (Rosenzweig et al., 2014; Asseng et al., 2013; Martre et al., 2015). This ensemble approach is cumbersome, usually computationally and financially expensive as it requires concerted efforts to collect large amounts of data and person-hours to calibrate and validate the models.

Projects such as the Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP) and its component Agricultural Model Intercomparison and Improvement Project (AgMIP) are international efforts to run an ensemble of global gridded crop models (GGCMs) with the same climate inputs and assumptions by different institutions across the globe (Rosenzweig et al., 2014; Warszawski et al., 2014). GGCMs are process-based models that simulate crop yield at a coarse spatial resolution (0.5^{o}x 0.5^{o}) (Elliott et al., 2013; Rosenzweig et al., 2014; Rosenzweig et al., 2013; Frieler et al., 2015) as a function of the global carbon, nitrogen, water and energy fluxes (Frieler et al., 2015). These models are useful in large-scale studies but often overlook finer spatial variability. This project provides a framework for comparison between different types of models and addressing uncertainties in crop yield modeling on a global scale and results in a wealth of simulated data for analysis on crop yield response to climate.

Statistical modeling is the less data-intensive approach to estimating crop yield response to climate conditions. Traditionally, these models use historical observational yield and climate data (hereinafter in this paper referred to as *observation models*). Observation models utilize cross-sectional or time series observations or a combination of both (also known as panel data) to examine impacts of climate on yield over time and space. Panel linear regression, which develops a functional relationship between yield and climate data using panel data, is an increasingly used method to study climate impacts on yield (Blanc and Schlenker, 2017; Roberts et al., 2017). In addition to being more efficient than process-based models, observation models can take advantage of the increasingly large datasets collected from actual crop fields, which encompasses a wider range of climate, environmental and management conditions of crop growth than the laboratory field experiments used in process-based models. Using these observations, the observation model can implicitly control for actual field conditions, including management

practices and farmer adaptation plans to shocks (Blanc and Schlenker, 2017; Lobell and Burke, 2010; Roberts et al., 2017). However, such a reduced-form of crop yield-climate relationship has several limitations, including: i) bias towards omitted variables (such as soil moisture, management practice, CO_2 fertilization impacts, etc.), ii) estimates of climate impacts on yield varying with different datasets, and iii) assumption that yield-climate relationship remain unchanged to the future (assumption of stationarity) (Schauberger et al., 2017; Tack et al., 2015). In climate change studies, yield projections of observation models as a function of future climate are usually constrained by past relationships between climate variables and yield, and thus need to be interpreted with caution (Blanc and Schlenker, 2017; Shi et al., 2013; Blanc, 2017b).

In recent years, another approach has emerged, called *emulators*, which is a hybrid between process-based and statistical models. Using the concept of 'perfect-model' introduced by Lobell and Burke (2010) - in which simulated yield outputs from a process-based model are considered to be the 'perfect', true observations - various studies have trained statistical models on simulated yield outputs. The emulator approach has been used to compare statistical and process-based models (Schlenker and Lobell, 2010; Roberts et al., 2017), to analyze physiological or management mechanisms that affect yield response to climate factors (Lobell et al., 2013; Ostberg et al., 2018; Schauberger et al., 2017) or to develop simple, efficient tools for predicting yield change under different climate scenarios (Blanc and Schlenker, 2017; Blanc and Sultan, 2015; Oyebamiji et al., 2015). An advantage of emulators is that they utilize the dynamic yield-climate relationships simulated by process-based models under climate change scenarios, while retaining the simplicity of statistical models, thus offering a less computationally expensive alternatives to process-based models (Oyebamiji et al., 2015).

The emulator approach has been explored by a few authors using yield outputs and climate forcings from the ISIMIP database (Rosenzweig et al., 2014). The ISIMIP database provides simulated datasets of estimated yield response to different climate change scenarios using combinations of GGCM-simulated data and their corresponding global circulation model (GCM) inputs. Blanc (2017b) and Oyebamiji et al. (2015) advocated emulators as useful and less computationally expensive than running process-based models, proving that they are capable of capturing the relationship between yield and climate change under a range of climate scenarios, radiation representative concentration pathways (RCPs) and management practices, as well as

predicting yield change to out-of-sample climate scenarios. Mistry et al. (2017) showed that the GGCMs in the ISIMIP– Fast Track database cannot reproduce the historical actual yield data at the county level, mostly because of the differences in the model inputs (climate data, fertilization, soil water availability, etc.) and the different sources of observations that the models were calibrated against. However, by constructing an ensemble of yield-climate functional relationships from the GGCM-simulated outputs, Schauberger et al. (2017) showed that the ensemble means resemble the historical relationship between yield and climate. Ruane et al. (2017) examined the strengths and weaknesses of using emulators based on process-based crop models. The study encourages the use of emulators as a more computationally-efficient method for assessment of agricultural sectors, and found that the GGCMs could provide comprehensive coverage for crop growth over large areas.

The use of statistical models, including both emulators and observation-based models, to study climate change impacts on crop production has been done at the global level (Blanc, 2017b; Blanc and Sultan, 2015), or for the conterminous US (Mistry et al., 2017) and the Corn Belt (Schauberger et al., 2017). However, there have been very few studies that focus specifically on the southeastern US, where there have been limited studies on the impact of climate change on crop production and water resources despite considerable agricultural activity and having experienced more droughts in recent years (Stocker, 2014). Emulators and observation models have been used separately in previous studies, and very few papers focus on analyzing the differences between these models (Schauberger et al., 2017; Roberts et al., 2017; Mistry et al., 2017). Therefore, this study aims at the following goals: i) estimating yield changes under future climate scenarios, ii) analyzing differences in how the emulators simulating yield response to climate compared to the observation model, and iii) highlighting role of using an ensemble for future work. The emulators were developed using simulated data generated by GGCMs' and their corresponding GCM inputs under RCP 8.5 from the ISIMIP-Fast Track database. The observation model were developed using historical yield from the USDA NASS database, and climate data from the University of Idaho Gridded Surface Meteorological Dataset (Abatzoglou, 2013). The study aims to provide more insight into modelling yield responses to climate change in a robust way.

5.2 *Methodology*

5.2.1 Data

5.2.1.1 Climate data

Historical climate data were taken from the University of Idaho Gridded Surface Meteorological Dataset (Abatzoglou, 2013). This dataset is a combination of two datasets: the temporally fine data from North American Land Data Assimilation System Phase 2 (NLDAS-2) (hourly time scales and 1/8th degree (approximately 12 km x 12 km) resolution, and the spatially rich data from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) (high spatial resolution (800 m)) climate surfaces of temperature, precipitation and dewpoint temperature at monthly timescale). ArcGIS was then used to aggregate climate data from the $1/8^{th}$ gridcell level to the county level by weight-averaging climate data of the grids in each county, with the weight being the faction of the county area that the grid occupies. The historical climate data time series covers the period of 1979-2016.

For training of the emulators, three GCMs selected as climate forcings for the GGCM-simulated yield data for this study were GFDL-ESM2M, NorESM1-M and HadGEM2-ES, under the representative concentration pathway (RCP) 8.5. These climate scenarios represent the low, medium and high, respectively, climate change in response to the radiative forcing pathways of 8.5 W/m^2 (Blanc, 2017b; Blanc and Sultan, 2015). The RCP 8.5 scenario was chosen to examine impacts of climate change under highest level of warming. The data are in the $0.5x0.5^o$ gridcell resolution and covers the period of 1950-2099. The same three GCM scenarios (GFDL-ESM2M, HadGEM2-ES and NorESM1-M) and same concentration pathway (RCP 8.5) were used as climate input into all the statistical models for yield change estimations. For this purpose, data of these three GCMs were taken from the University of Idaho Gridded Surface Meteorological Dataset (Abatzoglou, 2013) which offer better resolution as these data were further refined from the original CMIP5 dataset.

In the RCP 8.5 pathway, temperature is predicted to increase significantly by 1.4 – 2.2^oC by mid-century and 4 – 7^oC by the end of the century in all three climate scenarios. HadGEM2-ES projects the highest level of warming amongst the three GCMs (see Figure 5-1). Precipitation, however, shows no apparent trend in the future, except for decrease in the HadGEM2-ES for both the near future and far future periods (see Figure 5-2).

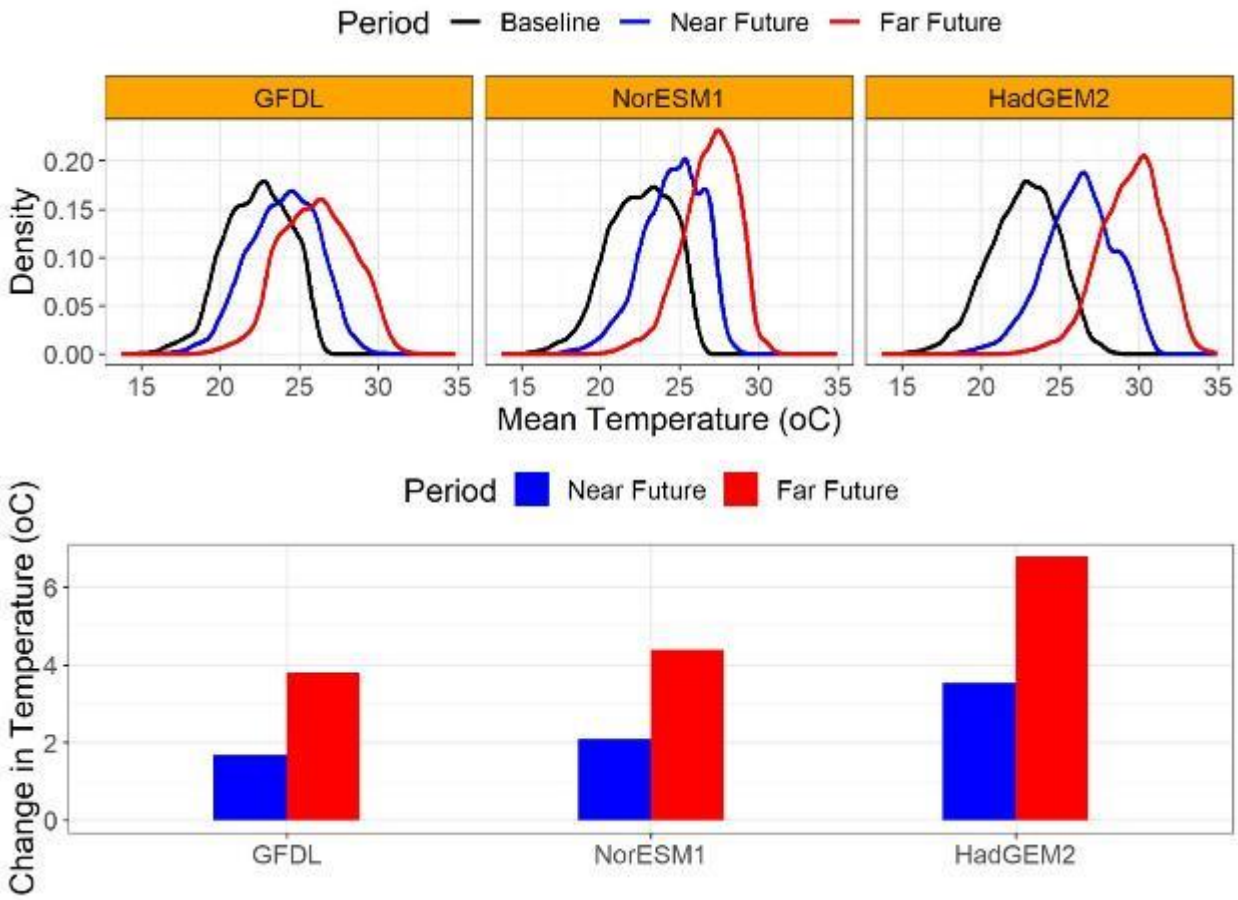

Figure 5-1 Changes in distribution (upper panel) and average (lower panel) of the growing season (April to September) daily mean temperature comparing the baseline (1980-2009) to the near future (2030-2059) and far future (2070-2099) for the three GCMs (GFDL-ESM2M, NorESM1-M and HadGEM2-ES)

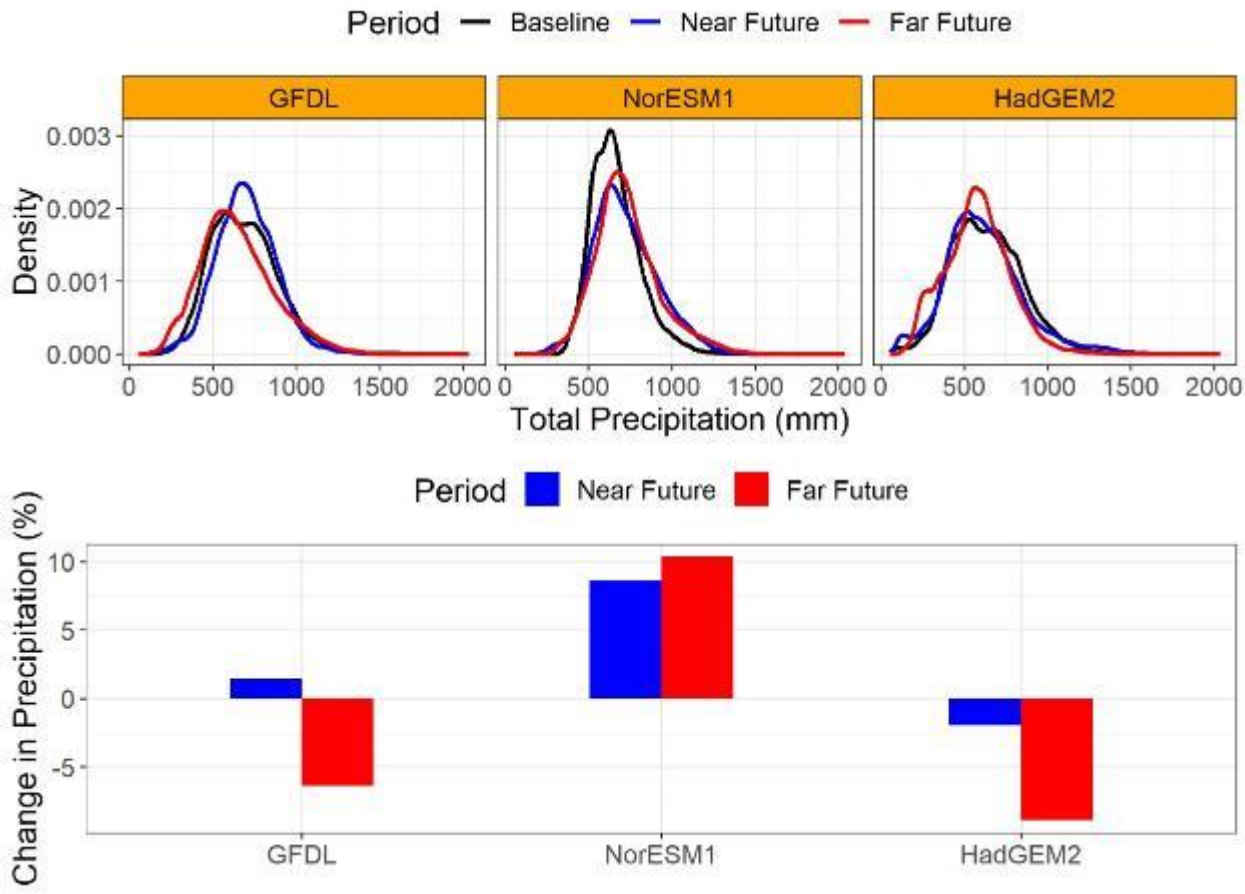

Figure 5-2 Changes in distribution (upper panel) and average (lower panel) of the growing season (April to September) precipitation comparing the baseline (1980-2009) to the near future (2030-2059) and far future (2070-2099) for the three GCMs (GFDL-ESM2M, NorESM1-M and HadGEM2-ES)

5.2.1.2 Actual and simulated yield observations

Historical yield observations in the US have been collected via annual surveys or 5-year census studies by the USDA National Agricultural Statistics Survey NASS) (https://quickstats.nass.usda.gov/). Yield observations for maize and soybean in the study region were collected for the period of 1979 and 2016. Counties selected for developing the observation model were those that: i) have more than 15 years in the record, and ii) cover a total of more than 90% of the total crop production area in the region.

Simulated maize and soybean yields were outputs from three GGCMs in the ISIMIP Fast Track: Lund Potsdam-Jena managed Land (LPJmL) (Sitch et al., 2003; Bondeau et al., 2007), the parallel Decision Support System for Agro-technology Transfer (pDSSAT) (Elliott et al., 2013; Jones et al., 2003) and the Predicting Ecosystem Goods And Services Using Scenarios (PEGASUS) (Deryng et al., 2011). These yield outputs were generated by these three GGCMs

with climate forcings from the three GCMs described in section *5.2.1*. Because maize and soybean in the region are mostly rainfed, and it is very difficult to evaluate CO_2 fertilization effect in the actual data (Lobell and Asseng, 2017), only yield outputs simulated under the rainfed and "no CO_2 fertilization" effect conditions were selected for this study. The selected yield simulations most closely approximate the conditions reflected in the historical data. The yield outputs and climate data obtained from ISIMIP are at the 0.5^ox0.5^o grid cell resolution and expand the period of 1979-2099. Grids that have more than 30 zero observations out of the 139 years were removed.

*Table 5-1 Summary of maize yield data during the historical period 1980 - 2004. Unit for yield statistics is in ton/ha (except for the Mann Kendall z-score). The *, ** and *** indicates the trend significance level at p < 0.05, 0.01 and 0.001, respectively*

a) Maize

GGCM	GCM	Max	Mean	Min	SD	Mann Kendall z-score
Observation	Historical	12.12	5.46	0.63	1.84	3.76**
LPJmL	GFDL	11.69	7.06	0.98	1.79	-0.91
	HadGEM2	17.55	6.49	0.31	2.15	-1.80
	NorESM1	10.07	7.28	2.49	1.13	1.99
pDSSAT	GFDL	13.06	6.45	0.07	1.43	-0.21
	HadGEM2	13.13	6.00	0.34	1.82	-1.33
	NorESM1	12.62	6.64	0.07	1.34	0.58
PEGASUS	GFDL	21.29	8.08	0.03	3.45	0.68
	HadGEM2	19.31	8.05	0.00	3.19	-3.43**
	NorESM1	23.92	9.76	0.05	2.90	-1.24

b) Soybean

GGCM	*GCM*	*Max*	*Mean*	*Min*	*SD*	*Mann Kendall z-score*
Observation	Historical	3.90	1.83	0.34	0.55	3.43**
LPJmL	GFDL	3.10	2.05	0.05	0.44	0.77
	HadGEM2	5.66	1.98	0.01	0.55	0.07
	NorESM1	3.12	2.22	0.13	0.32	4.60**
pDSSAT	GFDL	5.61	2.48	0.00	1.63	-0.26
	HadGEM2	5.13	2.46	0.00	1.54	-2.36
	NorESM1	4.92	3.09	0.05	1.15	0.07
PEGASUS	GFDL	7.28	2.18	0.00	0.78	1.56
	HadGEM2	5.70	1.98	0.00	0.84	-2.08*
	NorESM1	5.68	2.47	0.01	0.66	-1.19

In order to understand the differences between the observed yields and simulated yields, the mean, maximum, minimum and trend of yields during the 1980-2004 period for each of the yield datasets are summarized in Table 5-1. The non-parametric Mann-Kendall trend test was

applied to estimate the yield trend during this period for all datasets. The z-scores from the test was presented in Table 5-1 with the p-value for significance. The results show that the GGCMs often over- or under-estimate observed historical yield or the increasing trend of both maize and soybean during this period. The PEGASUS model estimated mean and maximum maize yields similar to the observed data, though with a larger standard deviation. The other two GGCMs overestimated mean and maximum maize yields much larger than the observations. All of the models simulated mean soybean yields fairly close to the observation yields during the same period, especially pDSSAT estimated much wider yield range for soybean. Almost all GGCMs were unable to capture the increasing trend observed over time when using Mann-Kendall method. The GGCMs were simulated under constant management practices and technological conditions, which are typically the driving factors of the increasing yield trend in reality. Because of difference in yield simulations by the GGCMs and observation yield, the study focused on comparing relative yield changes predicted by the statistical models rather than the absolute yield results.

5.3 *Statistical crop yield modeling*

5.3.1 Model specification

Yield response to climate is non-linear: as climate factors increase, yield increases to a certain threshold before declining (Schlenker and Roberts, 2009). This non-linearity can be captured reasonably well by annual mean temperature and total precipitation of the growing season and their quadratic terms. The model specification is as follows:

$$log(y_{ct}) = \alpha_c + \beta_1^T * T_{c,t} + \beta_2^T * T_{c,t}^2 + \beta_1^P * P_{c,t} + \beta_2^P * P_{c,t}^2 + f(t) + \mathcal{E}_{c,t} \quad \text{Eq. 5-1}$$

where: $log(y_{ct})$ is the natural logarithm transformation of yield (ton/ha), T is the mean temperature of the growing season (^oC), P is growing season total rainfall (mm), $f(t)$ controls the effect of time-variant factors in the models; β_i are the coefficient estimates of climate variables, α_c is the time-invariant county fixed effect and $\mathcal{E}$ is the error term. The characters c and t indicate the county and year.

The growing season was fixed from 1st April to 31st September in this study, which is the typical growing season period for the study area. While each GGCM model has a unique internal process of choosing the planting and maturity dates, fixed growing season were used because: i) climate variables derived from the model-specific growing season did not improve the model nor drastically change the relationship between yield and climate variables (Schauberger et al., 2017),

and ii) a fixed growing season helps to eliminate the endogenous influences of the crop models on the statistical model outcomes (Mistry et al., 2017).

One of major difference between the GGCM-simulated and historical yields in this study is the range of the climate inputs. Statistical models based on historical data assumes that past relationships will continue in the future, therefore the ability to predict yield response to high temperatures that never occurred before may be limited. Emulators takes advantage of the effort to mathematically capture biophysical interactions of yield and future climate within the GGCM physiological environment, thus may provide some insight into how yield may behave to unprecedented climate events. Therefore, while the training datasets for the observation model were limited to year 2016, the period for training data of the emulators were 1979-2099.

5.3.2 Panel linear regression

Yield and climate data were paired together to form an unbalanced panel dataset for maize and soybean. To develop an emulator, a panel was constructed for each GGCM and each crop, consisting of yield outputs and corresponding climate forcings from all three GCMs. Using all three GCMs gives a wider range of climate data, enhancing the emulators capability of explaining yield variation by climate factors. For the observation model, actual yield and climate data were used, while simulated yield and climate data were used for building the emulators, which are referred to by the GGCM used to estimate yield followed by "_m", e.g., LPJmL_m.

A panel linear regression with fixed effects is fitted through each panel; the fixed effects are to control time-invariant factors that are not included in the model specification but may affect yield, such as soil type or location characteristics (Brüderl and Ludwig, 2015). For the observation model, the trend term $f(t) = \lambda_s * t$ represents the yield trend (λ_s) throughout the period 1979-2016 as there is a significant positive trend driven by advances in technology and management practices over this period. For the emulators, however, technological and management inputs were held constant throughout the simulation run, therefore using time dummies that represents the year-specific influence on crop yield would be more appropriate (Mistry et al., 2017). To account for heteroscedasiticity in the model residuals that are caused mostly by spatial autocorrelation (temporal correlation does exist though much weaker (Schlenker and Lobell, 2010)), robust clustered standard error at the cross-sectional unit (a county for the observation and a grid for the

emulators) is used to determine whether the model coefficients are statistically significant (Cameron et al., 2011).

5.3.3 Estimating yield change

Yield changes are expressed as relative changes from the near future period (2030-2059) and far future period (2070-2099) to the baseline period (1980-2009). The changes, Δy, is calculated as follows:

$$\Delta y = \frac{y_{future} - y_{baseline}}{y_{baseline}} = \frac{\exp\left(f(climate_{future})\right) - \exp\left(f(climate_{baseline})\right)}{\exp\left(f(climate_{baseline})\right)} \quad \textit{Eq. 5-2}$$

where: $climate_{future}$ and $climate_{baseline}$ are climate data taken from the near/far future period and the baseline period of each GCM accordingly.

5.4 *Results*

5.4.1 Performances of statistical models

Coefficient estimates for each climate variable and their squared terms, as well as the model coefficient of determination (R^2), are shown in Table 5-2. The R^2 values show how much variation in the original crop yields could be explained by the models. Most of the statistical models result in fair to good R^2 values in the range of 0.34 to 0.62 for maize and 0.30 to 0.71 for soybean. The observation model generally captures the observed yield variation well for a biological process (0.62 for both maize and soybean). Amongst all models, LPJmL_m and pDSSAT_m both have the highest R^2 values for maize (0.68) and soybean (0.71), indicating that the emulators were able to capture the majority of yield variation produced by the respective GGCMs. For both crops, PEGASUS_m has the lowest R^2 values (0.34 for maize and 0.30 for soybean), which can be attributed to the dynamic planting date selection mechanism of the model (Rosenzweig et al., 2014). Many of planting and mature dates selected by PEGASUS fell outside the historical growing season (fixed from April 1st to September 30th in this study). In contrast, LPJmL and pDSSAT assumed a general fixed planting date in April for each grid cell throughout the simulation runtime (Rosenzweig et al., 2014) (see summary of GGCMs in Appendix B).

Table 5-2 Estimate coefficients and R-squared for the observation model and emulators

a) Maize

Model	T	T-squared	P	P-squared	R-squared
Observation	1.8e-01	-7.1e-03	2.1e-03	-1.2e-06	0.62
LPJmL-m	-4.2e-02	-2.7e-04	4.0e-03	-1.9e-06	0.68
pDSSAT-m	4.7e-03	-1.6e-03	4.4e-03	-2.5e-06	0.50
PEGASUS-m	3.2e-02	-3.3e-03	2.5e-03	-1.6e-06	0.34

b) Soybean

Model	T	T-squared	P	P-squared	R-squared
Observation	2.3e-01	-6.6e-03	3.3e-03	-2.0e-06	0.62
LPJmL-m	-3.7e-02	7.8e-05	3.7e-03	-2.0e-06	0.43
pDSSAT-m	7.8e-02	-1.0e-02	1.1e-02	-5.1e-06	0.71
PEGASUS-m	-1.5e-01	-1.1e-03	2.7e-03	-1.7e-06	0.30

5.4.2 Yield sensitivity to climate factors

In order to gain deeper insight into model performance, a sensitivity analysis of yield response to changes in temperature and precipitation for each model is shown in Figures 5-3 and 5-4, respectively. All models show broadly similar non-linear response of crop yield to changes in temperature and precipitation. Higher growing season mean temperature and reduced total precipitation result in negative impacts on yields, while increase in total precipitation would only increase yield to a certain level before excessive precipitation results in yield decrease, though the point after which yield declines varies across models. Figures 5-3 and 5-4 show that the observation model shows higher sensitivity of maize yield to temperature than all three emulators. For soybean, the observation model shows moderate yield reductions in comparison to two of the three emulators. For both crops, LPJmL_m estimates the least yield response to temperature, while pDSSAT_m shows highest level of sensitivity to temperature for soybeans. All models showed negative impacts of reduced precipitation on crop yield, with pDSSAT_m being the most sensitivity to decreases in precipitation for both crops. The sensitivity analysis indicates that all models agree that the impacts of high temperature and reduced precipitation would severely penalize crop production.

The differences in yield response to temperature and precipitation between the emulators and the observation models and amongst the emulators result from the underlying assumptions and

model structures of the original GGCMs, especially in terms of interactions between climate and crop growth, and stress mechanisms. Compared to the gradual increase in historical fertilization application, the GGCMs either assume constant level of nitrogen, irrigation and other management practices (LPJmL), dynamic application with fertilization rates observed in the year 2000 (pDSSAT) or annual application with International Fertilization Association (IFA) NPK application rate (PEGASUS) (Rosenzweig et al., 2014), while in reality farmers can adjust level of inputs to fit the needs of the crops, thus result in lower yield sensitivity to climate factors than some models. Amongst the emulators, stress mechanisms may determine biggest differences in yield sensitivity to crops. LPJmL does not explicitly simulate N, thus yield simulations of this model tend to be more optimistic than pDSSAT and PEGASUS, which has explicit simulation of N (Rosenzweig et al., 2014). The higher sensitivity to temperature and precipitation of nitrogen stress-explicit models indicates with the current level of crop management and technology, impacts of climate change on yield (especially extreme events such as heat or droughts), which agrees with indings in Rosenzweig et al. (2014).

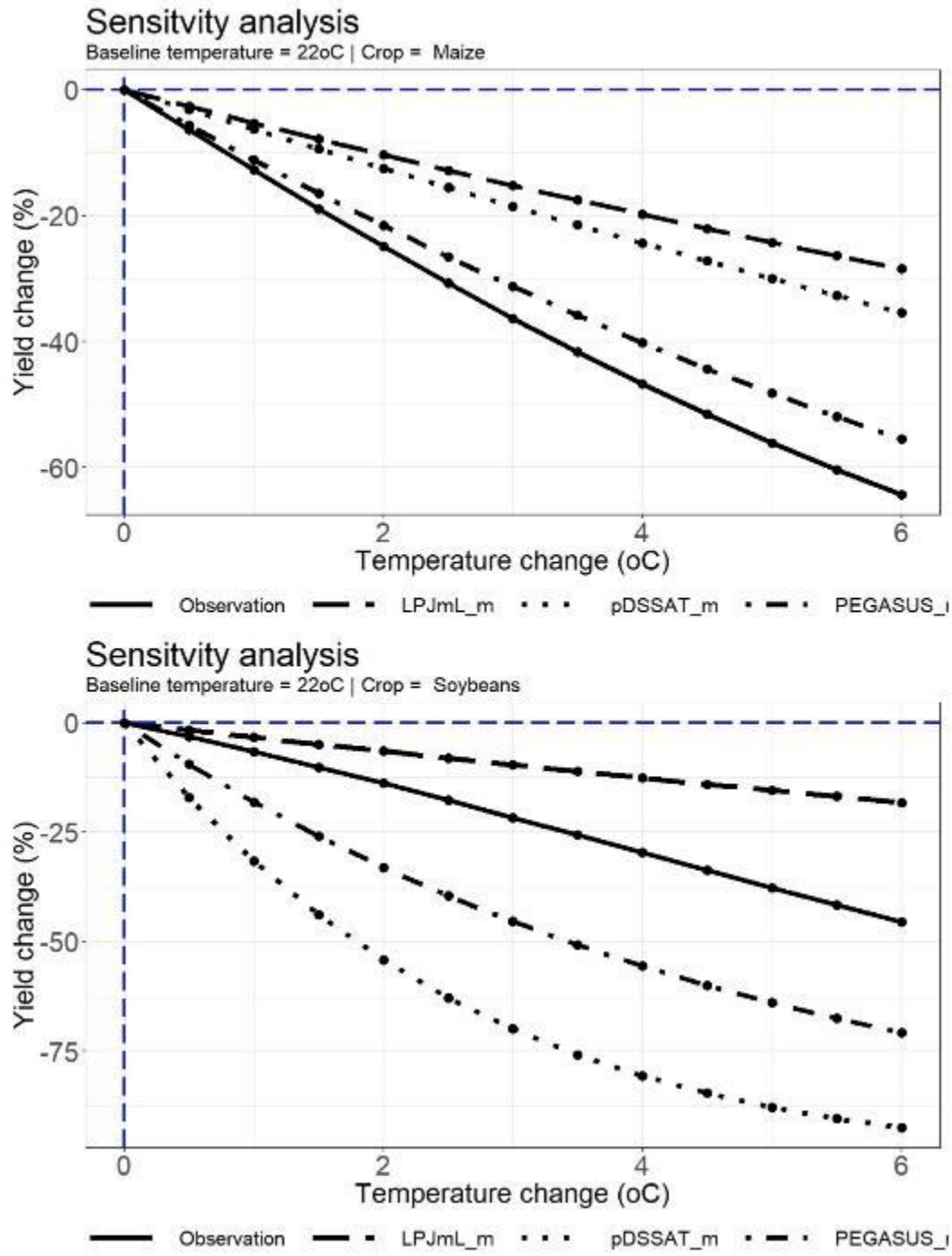

Figure 5-3 Sensitivity analysis of yield response to temperature change for maize (upper) and soybean (lower). Yield change is expressed as percentage of yield when the growing season mean temperature increases from the baseline (22ºC).

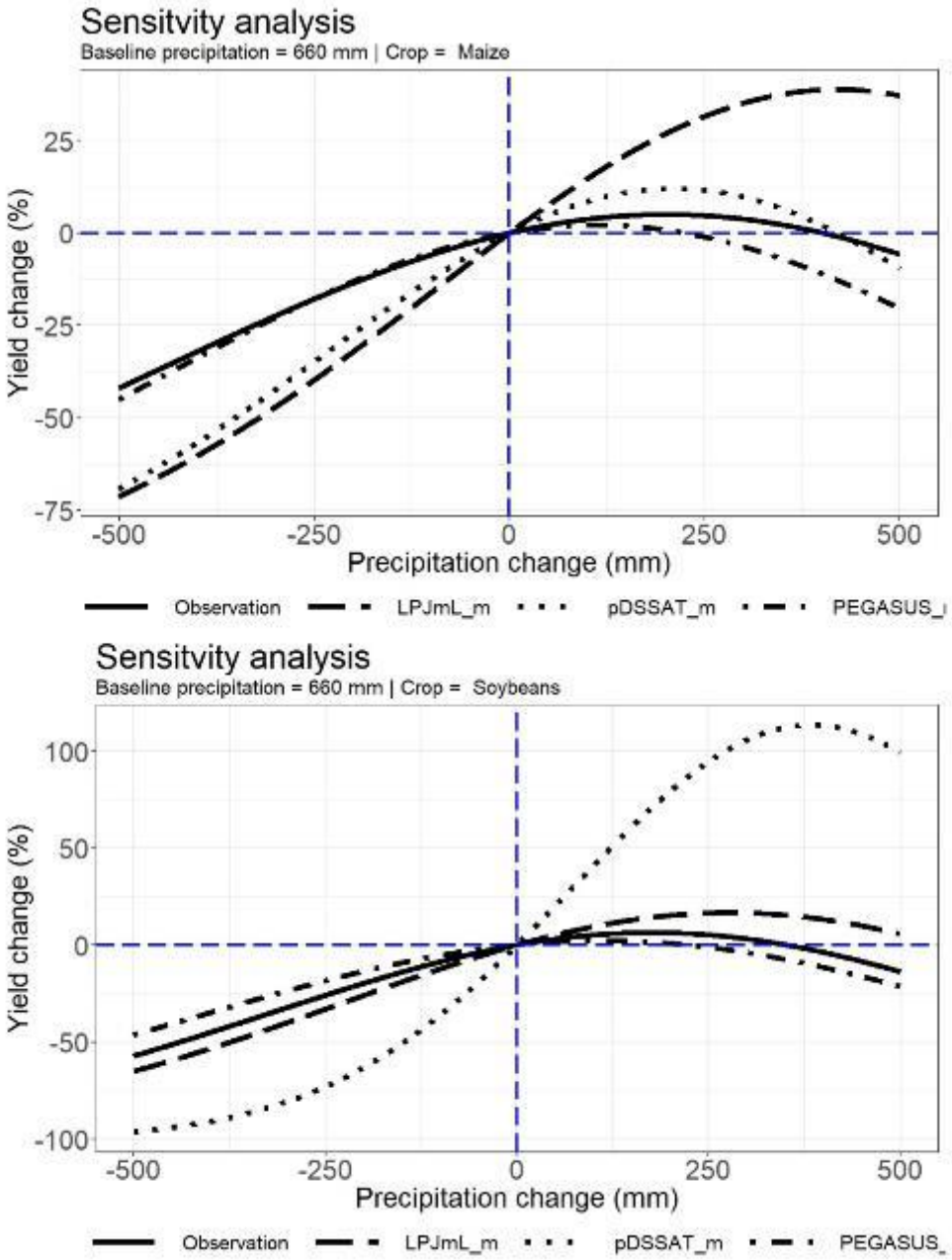

Figure 5-4 Sensitivity analysis for yield response to precipitation change maize (upper) and soybean (lower). Yield change is expressed as percentage of yield when the total growing season precipitation increases or decreases from the baseline (660mm).

5.4.3 Estimations of yield change under a high warming scenario

Yield response is most severely negative when predicted using HadGEM2-ES data. Under RCP 8.5, HadGEM2-ES predicts the largest increase in average temperature of 3ºC in the near future to 6.7ºC in the far future (Figure 5-1) and the largest average negative percent changes in cumulative growing season precipitation in the near (2.1%) and far future (7.7%) (Figure 5-2). Yield response estimated is less severe using GFDL-ESM2M and NorESM1-M. Across all three

GCMs, both the observation model and the emulators project decline in median maize yields of the study area (Figure 5-5). The observation model estimates the largest reduction in maize yield under all scenarios, with a median decline of 20%, 25% and 46% by the near future century (2030-2059) and 46%, 51% and 73% by the far future (2070-2099), under the GFDL-ESM2M, NorESM1-M and HadGEM2-ES, respectively. Under the respective climate scenarios, LPJmL_m predicts the least yield changes, with median reduction of 6%, 4% and 22% by the near future, and 25%, 16% and 38% by the far future (for more details, see Appendix B).

Soybean yield is predicted to decline across all models, though the results vary more widely between crop models and GCMs than in the case of maize. Similar to maize, results from LPJmL-m predict the lowest yield changes across the study area (median reduction of 3%, 4% and 15% by the near future, and 17%, 11% and 26% by the far future, under the GFDL-ESM2M, NorESM1-M and HadGEM2-ES, respectively). And, the model predicts small increases in yield in some counties in the near future under the GFDL-ESM2M and NorESM1-M scenarios. pDSSAT-m predicts a drastic reduction in soybean yield (median decline of 27%, 35% and 53% in the near future, and 82%, 81% and 96% in the far future, under the GFDL-ESM2M, NorESM1-M and HadGEM2-ES, respectively).

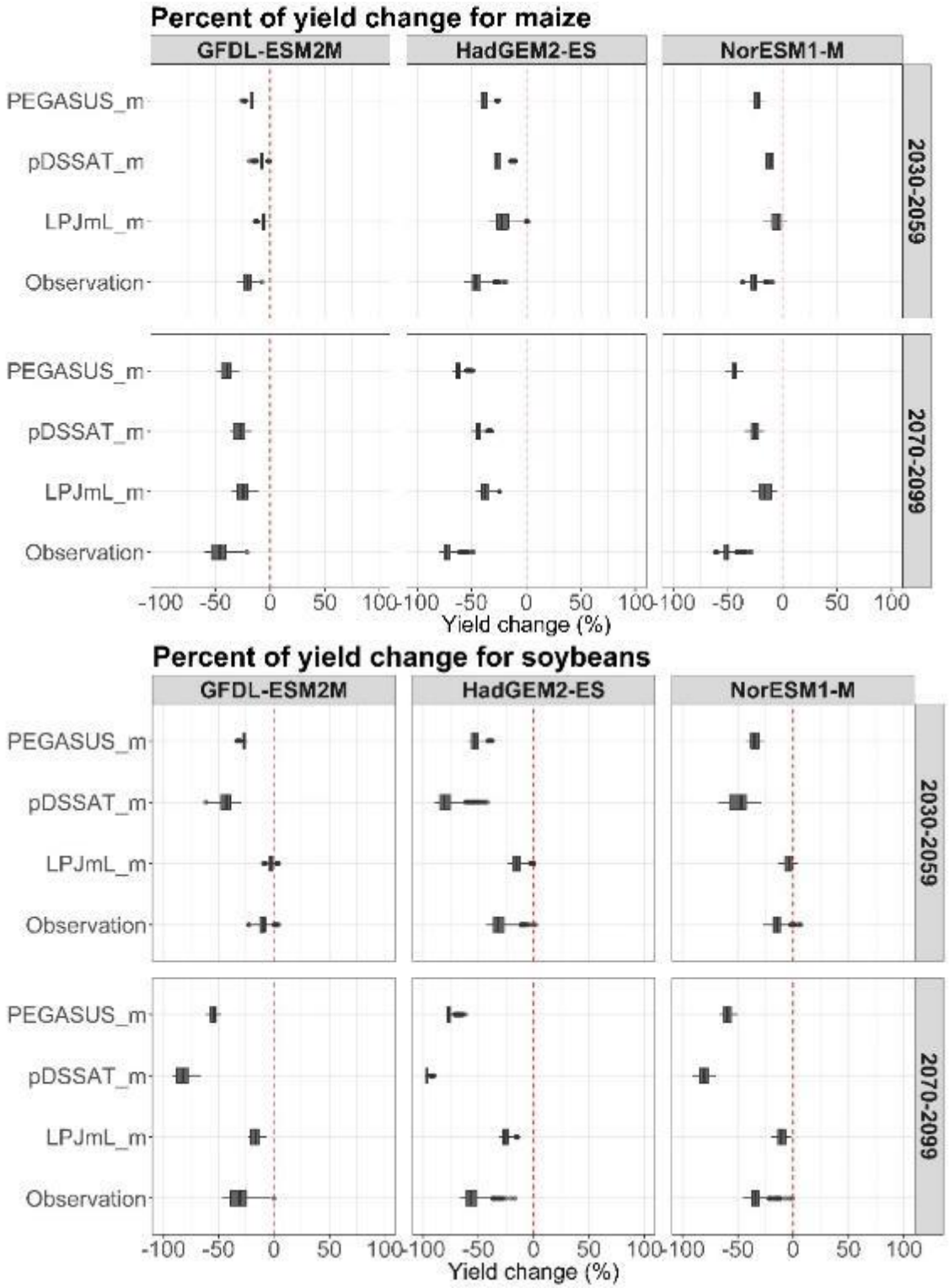

Figure 5-5 Predictions of yield changes in maize (upper) and soybean (lower) yield according to three different GCMs by observation models and emulators. The boxplots show yield change predictions throughout the study area of the mid-century (2030-3059), and end of century (2070-2099) to the baseline period (1980-2009).

While all models predict a decline in median yield under future climate scenarios, climate change impacts on yield are not uniformly distributed as shown by the boxplots in Figure 5-5. One of the major reasons for the differences in magnitude of yield change amongst the counties in the study area is the historical growing season mean temperature. Various studies show that regions that are historically warmer may experience more yield decline than regions with cooler climate, as the mean temperature of the growing season is already closer to the optimal threshold (Rosenzweig et al., 2014; Schauberger et al., 2017). In order to how yield change in relation to average historical growing season mean temperature, the counties in the study area clustered into three homogenous groups by their historical growing season mean temperature (1979-2016) of 19^{o}C, 22^{o}C and 24^{o}C (Figure 5-6). Then, maize and soybean county yield changes estimated by each model are plotted against its temperature change predicted by the corresponding GCM for the near and far future (Figures 5-7 and 5-8). These figures show that all models predict that the historically warm counties (group 3) experience the highest estimated yield reductions, while group 1 generally has lowest yield decline under all three climate scenarios since the historical temperature of this group is lower. There is generally a strong linear correlation between yield changes and temperature increases. As crops in the study area are mostly rainfed, this indicates that the with increase in temperature, there is a need for increase in irrigation for crops. For mapping of spatial distribution of yield changes, see Supplemental Information.

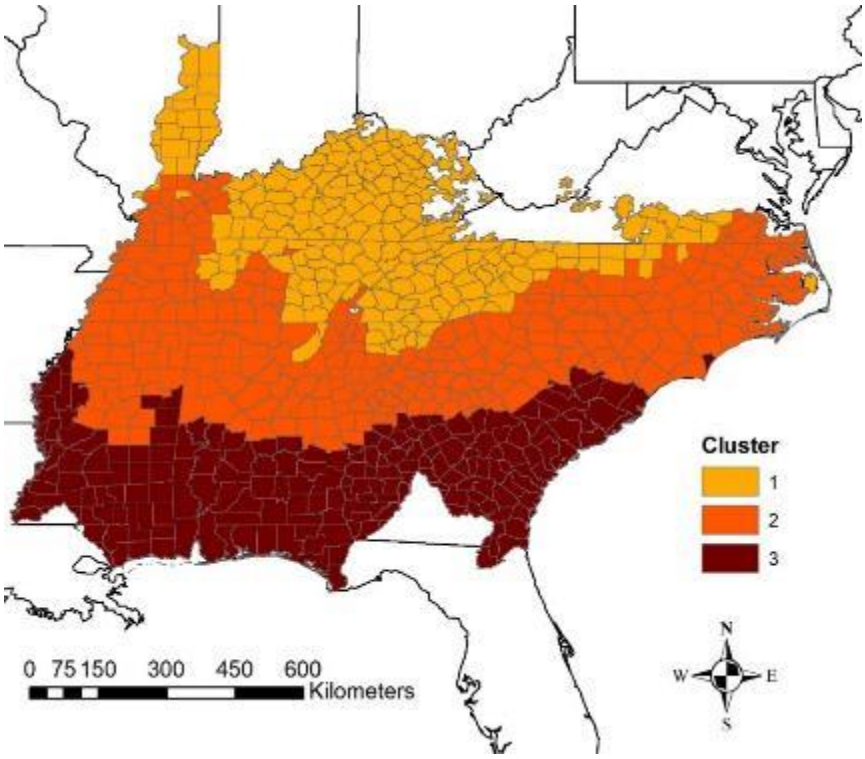

Figure 5-6 Spatial distribution of growing season mean temperature in the study region based on historical data (1979-2016). Cluster 1 is characterized by a 30-year annual mean growing season temperature of 19°C; Cluster 2 is 22°C and Cluster 3 is 24°C.

5.5 *Discussion*

By using the emulators and the observation model to examine yield changes under different future climate scenarios, the study shows that overall, median yield of the study area will decline as temperature rises if there is no change in adaptation and management plan, more severely for historically warm areas that experience high level of warming coupled with low level of precipitation. The emulators and observation model generally capture the nonlinearity of yield response to precipitation and temperature, as shown in the sensitivity analysis. However, the emulators and observation models differ in their predictions of severity in yield reduction, because of difference in their sensitivity to temperature and precipitation changes. As mentioned above, the differences stem from the differences of the underlying assumptions on crop phenology, management practice, technology, environment and stress simulation amongst the GGCMs and between these assumptions to the historical reality.

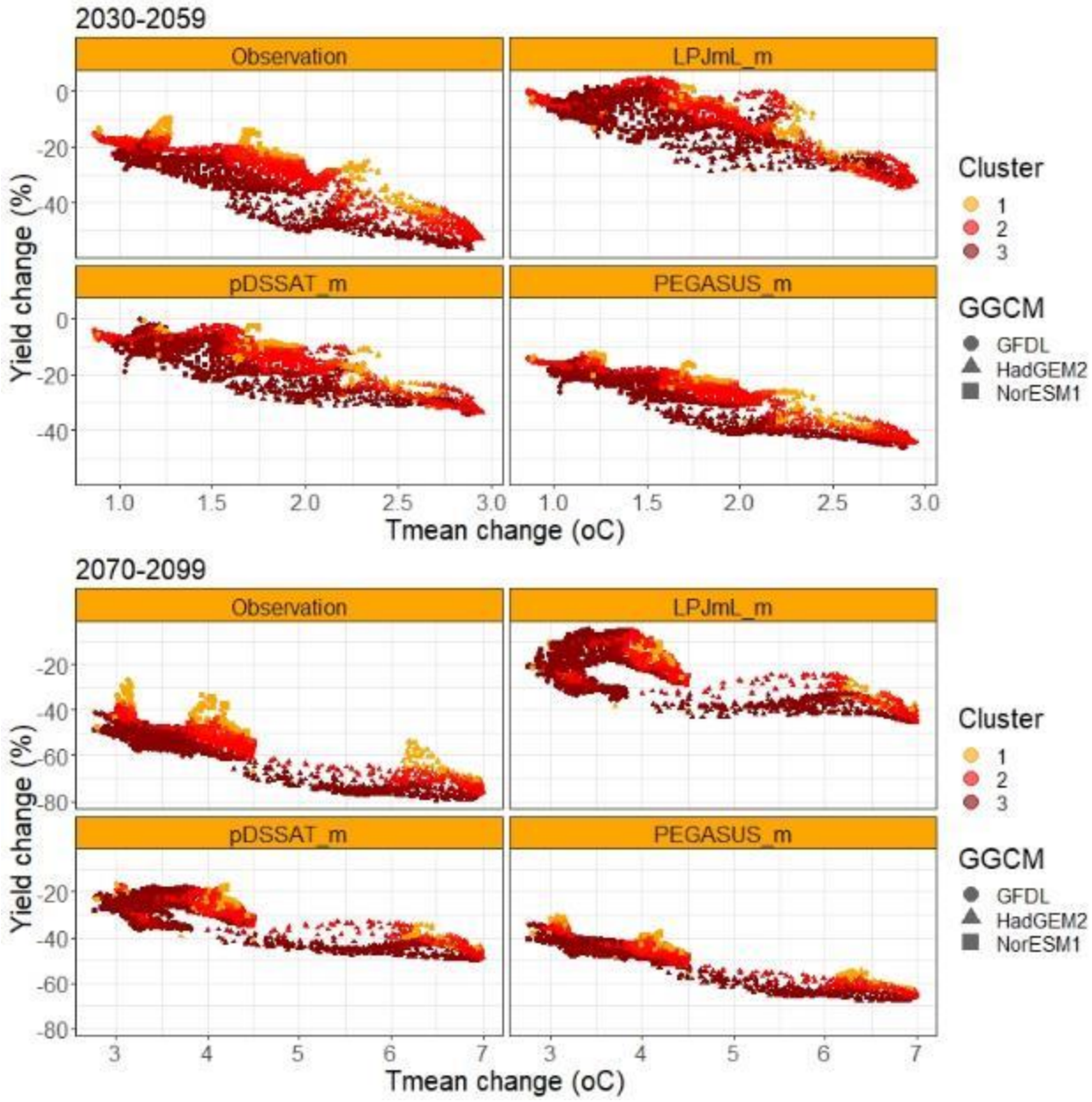

Figure 5-7 Yield responses to changes in temperature under the three different GCMs for maize for the near future (upper) and far future (lower). The three clusters of counties by temperature are presented by the color (1 – orange, 2 – red and 3 – dark red). GCMs are represented by circle (GFDL-ESM2M), triangle (HadGEM2-ES) and square (NorESM1-M).

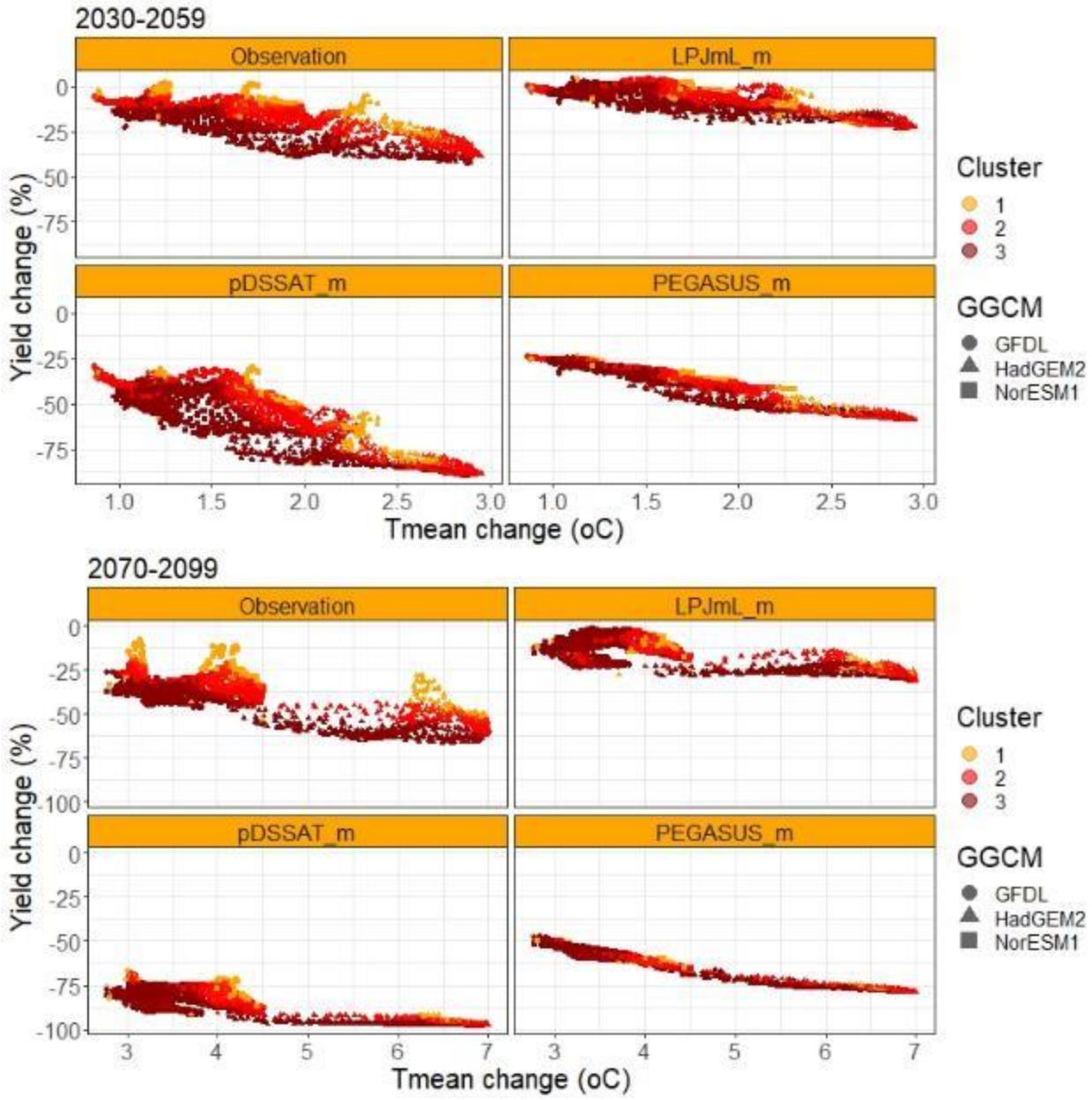

Figure 5-8 Yield responses to changes in temperature under the three different GCMs for soybean for the near future (upper) and far future (lower). The three clusters of counties by temperature are presented by the color (1 – orange, 2 – red and 3 – dark red). GCMs are represented by circle (GFDL-ESM2M), triangle (HadGEM2-ES) and square (NorESM1-M).

As observation models are built on historical data, with climate data limited to the historical range, estimations of yield change by the observation model reflect the historical interactions between climate factors and crop growth processes. Compared to emulators, observation models are better at capturing the actual conditions of a region, thus help to provide more "ground-truth" information regarding crop production. However, since the level of warming and precipitation projected by the GCMs into the future have not yet occurred in the historical record, the extrapolation of the historical relationships of climate and yield may not reflect the realistic

response to future climate. On the contrary, the GGCMs simulate yield response to climate using phenological and biophysical principles, they are theoretically capable of exploring crop yield formation in a new environment that has never occurred before. Outputs from GGCMs also reflect the pure relationship between climate and crop yields, as other variables are better controlled, as opposed to historical observations where relevant factors may be unavailable. However, since the GGCMs differ from each other in terms of underlying assumptions and equations used to represent the environment in which crops are grown, there is a risk of model bias in these model outputs. Since outputs from GGCMs may not coincide with the historical observations of a location, it is usually used to predict relative yield changes between the future and the baseline period, rather than the absolute yield. Emulators will inevitably inherit all the advantages and disadvantages of the original GCMs. In general, observational models are best to produce a true historical relationship and project absolute yield value in the future, while emulators may be best to analyze the relationship between yield and climate under a controlled environment. With their advantages and disadvantages, it is strongly recommended to use of an ensemble of models is an approach to deliver quantify the uncertainties in climate change impacts on crop yields.

There are several aspects of yield response to climate change that was not addressed in the study. First, the use of growing season mean temperature and total precipitation for a fixed growing season is simple and efficient, but could be limiting. Especially when models have dynamic mechanism for selecting planting date, then the fixed growing season will reduce the significance of the relationship between yield and the climate variables. Second, the models do not examine the effects of carbon feralization (CFE) or impacts of pest or disease under future climate conditions, as it is still difficult to discern this effect from management practice and technological development using historical data. Most of experiments on impacts of CFE on crop radiation use efficiency are limited to low level of CO_2 concentration (approximately 550ppm), thus available data on CFE at high level of CO_2 are limited. Finally, this study could benefit from using more data from other GGCMs and GCMs. The simulated data in this study was limited to the ISIMIP Fast Track, which consists of only 7 GGCMs and 5 GCMs.

5.6 *Conclusions*

In order to have effective plans for future crop production in the southeastern US, an ensemble of statistical models was developed for climate change impacts on maize and soybean

yields by county in the near and far future under RCP 8.5. Future climate change is likely to impact crop production due to increases in temperature and uncertain rainfall pattern throughout the growing season, thus threatening food security. Results show that overall, all models predict that the study area median yield would decline. Median decrease in maize yield is estimated by 5% to 45% by the near future (2030-2059) and 16% to 72% by the far future (2070-2099), while median soybean yield is predicted to decline by 4% to 78% by the near future and 10% to 98% for by the far future for soybean. Reduction in yield is particularly critical in counties that are historically warmer (i.e., the southernmost counties). In general, the emulators are capable of capturing the non-linear impacts of climate change on yield and the direction of change similar to the observation model. Differences in the magnitude of changes come from the underlying assumptions for crop phenological processes and interactions between crop growth and limiting factors such as water, soil carbon and fertilizer. However, the use of the observation model and the emulators provide more insight into the analysis of climate change impacts on yield in the study area. With the increasing quantity of available data, using an ensemble of emulators could be an efficient and effective way for improving understanding and addressing uncertainty in crop production under future climate.

Chapter 6 Conclusions

6.1 Goals and objectives

Despite the tremendous advancement in technology and management practices, agriculture is still highly vulnerable to climate change (Zhongming et al., 2019). Agricultural production, in particular crop production, and food security are already being affected and will continue to face negative consequences of climate change at the global and regional scales. To manage and mitigate the risks of climate change to agriculture, long-term, large-scale and timely strategies play a crucial part. Therefore, agricultural managers, policy-makers and stakeholders must understand the extent of climate change impacts on agricultural productivity to develop effective, efficient management and adaptation plans. With the urgency of climate change threats, there exists a need for robust and timely modeling of climate change impacts on agriculture to assess future risks and corresponding adaptation strategies.

Climate change impacts on crop production are multifold but two aspects remain top priority for researchers: crop productivity and crop water demand. These two aspects need to be examined, taking into consideration the uncertainties in climate change projections and crop modeling approaches. This research answered key questions of great concern amongst scientists and decision-makers, including: *i) can a future warming trend during the growing season that may pose significant threats to crop growth be detected; ii) what is the direction and magnitude of crop yield change to temperature increase and precipitation pattern changes under future climate scenarios; and iii) how does crop net irrigation requirement change under future climate*. While this study does not focus on examining the non-climatic factors that may affect crop production, findings from this study can provide useful information for further study on areas that may experience significant changes in climate and crop production response to those changes, thus allowing strategic planning and allocation of resources to those areas.

6.2 Study findings

6.2.1 Finding 1

Significant trend in mean temperature detected across CONUS under future climate

To improve the level of confidence in assessing impacts of climate change on the environment and socioeconomic sectors, internal climate variability (ICV) needs to be quantified

and removed from climate change projections. By developing a statistical model using a high-resolution mean temperature dataset by Livneh et al. (2015) and a multi-model ensemble of statistically downscaled CMIP5 GCMs for RCP 8.5 to identify the signal to noise ratio to assess the strength of mean temperature signal in the CONUS, the study showed that that historically (1961-2010), CONUS-mean increase of temperature was approximately 0.5^{o}C, while under RCP 8.5 scenario, CONUS-mean temperature will increase by a 1.27^{o}C increase in the future (2011-2060). The SNR estimated from the large-scale observational ensemble (OBS-MACA), developed based on showed the historical increasing trend in mean temperature was only significant to some part in the CONUS (mostly southwester, some part of southeastern and northeastern region), with the critical months for maize and soybean production (June, July and August). However, the increasing trend will become more prominent in most part in the CONUS for the summer time (SNR between 4-6), where a CONUS-average temperature might increase by 1.52^{o}C. This poses risks of drought and heat exposure for crop production in lower latitude regions. Mean temperature in winter months for the western regions may also experience significant increase by over 1^{o}C, which might lead to decrease in snow pack in the spring.

The signal-to-noise ratio map (Figure 2-11) reveals whether the temperature trend is significant or not. For the past period, the winter time for most of the CONUS did not show a strong signal (SNR < 2), except for the southwestern (SNR between 2-4 for October, November and December). The summer months for the southwestern and some parts of the southeastern area, on the other hand, had high SNR, indicating the rising trend in temperature during this season in those areas were significant. In particular, June, July and August - the three critical months during the growing season for maize and soybean - show strong SNR (4-6), indicating there has been a significant increase in mean temperature for those months during the historical period. This summer time rising temperature trend seems to continue in the future period (2011-2060), as indicated by high SNRs for the most part in the CONUS. The rising trend in winter temperature also tends to become more prominent, especially for the western U.S.

6.2.2 Finding 2

Net irrigation requirement in rainfed crop areas may experience changes depending on crop type, growing season length, effects of elevated [CO_2] and planting date.

To ensure crop production to operate at the normal level, decision-makers must know if there is a risk of water shortage for crops, especially the historically rainfed ones, at what location for investments in infrastructure. The initial results show that there is a significant negative correlation between yield loss and droughts in most counties. Yield losses (40% on average for maize and 28% for soybean) are significant when abnormally dry conditions occur. Maize is more susceptible to drought than soybean. Particularly, droughts are responsible for most of yield loss in rainfed area.

Crop potential evapotranspiration and net irrigation requirement estimations from an ensemble of 18 downscaled GCM that show that there can be a wide range of projected changes in ET_c and NIR as compared to the historical baseline (1980 to 2009). ET_c and NIR estimations follow similar seasonal pattern mostly because changes in effective precipitation in the region are not large enough to offset the changes in ET_c. These changes depend on crop type, CFE assumptions and planting date scenario. The study finds that NIR tends to decrease under the fixed planting date scenario compared to the historical baseline due to a shortened growing season, and increase under the adaptive planting date scenario as the climatic demand for evapotranspiration increases due to higher temperature while the risk of accelerated maturity is reduced. The effects of elevated CO_2 level may reduce crop NIR under the same planting date scenario and RCP. Under the same planting date scenario and RCP, if CO_2 level is fixed at the contemporary level, NIR of maize will increase by 5.82 and 3.33 % in the near future and 4.18 and 1.37 % under RCP 4.5 and RCP 8.5, respectively, while NIR of soybean will increase by 13.53 and 4.51 % in the near future, and 9.10 and 2.05 % in the far future under RCP 4.5 and RCP 8.5, respectively. Crop growing cycle is altered by changes in temperature, and generally crop growing season will be shortened if no adaptation is made for the planting date. For maize, under the fixed planting date, the growing season length tends to be 3 - 15 days shorter in the future periods under RCP 4.5 (RCP 8.5); for soybean, the corresponding lengths are reduced by 7 (12) and 15 (20). If crops are adapted to the new climate, growing season generally starts over 5 (10) days earlier than fixed planting date in both the near and far future for maize and soybean under RCP 4.5 (RCP 8.5). Fixed planting date increases the risk of heat exposure, hastens crop maturity, which will lead to an overall reduction in biomass accumulation and yield. Generally, shorter growing season leads to lower ET_c, except for maize in the near future under RCP45, as the slight shorter growing season does not compensate for increase climatic evaporative demand from increased temperature. With adaptive planting date,

the risk of shortened growing season is reduced, thus the water requirement is enhanced under this scenario compared to a fixed date. Under the same planting date and RCP scenario, the effects of elevated CO_2 level may reduce crop NIR.

As about 75% of maize and 64% of soybean are rainfed in the region, the increase in drought under warmer climate will lead to more water needs for crops. Spatial pattern of NIR shows that the southern area may experience little change or reduction in NIR due to increase in crop water demand under all planting date or NIR scenario as the impacts of increased temperature on shortened growing season length may overcome or cancel out enhanced climatic evaporative demand. Of the crop-growing areas, the northwestern area near the Corn Belt, where crops were traditionally rainfed, may require investments in irrigation infrastructure, as this area may experience increase in NIR for both maize and soybean.

6.2.3 Finding 3

Under high level of warming, crop production may experience yield loss, without CFE and adaptation planning

Future climate change is likely to impact crop production due to increases in temperature and uncertain rainfall pattern throughout the growing season, thus threatening crop yield. For the purpose of estimating crop yield changes under climate change, this study employed a statistical modeling approach, which is more time-effective and less expensive than process-based modelling. To address the "assumption of stationarity" in statistical modeling, this study employed a new approach, called *emulators*, which produce a reduced-form relationship between crop growth and climate variables of the original process-based model, based on climate and crop model outputs from AgMIP. Results show that overall, all models predict that the study area median yield would decline if there is not adaptation to future climate. Median decrease in maize yield is estimated by 5% to 45% in the near future (2030-2059) and 16% to 72% in the far future (2070-2099), while median soybean yield is predicted to decline by 4% to 78% by the near future and 10% to 98% for by the far future. Reduction in yield is particularly critical in counties that are historically warmer (i.e., the southernmost counties). In general, the emulators are capable of capturing the non-linear impacts of climate change on yield and the direction of change similar to the observation model. Differences in the magnitude of changes come from the underlying assumptions for crop phenological processes and interactions between crop growth and limiting

factors such as water, soil carbon and fertilizer. However, the use of the observation model and the emulators provide more insight into the analysis of climate change impacts on yield in the study area. With the increasing quantity of available data, using an ensemble of emulators could be an efficient and effective way for improving understanding and addressing uncertainty in crop production under future climate.

6.3 Limitations and further studies

There are several aspects of crop yield and crop water demand response to climate change that were not addressed in the study. First, the use of growing season mean temperature and total precipitation for a fixed growing season in crop yield modeling is simple and efficient, but could be limiting. This is especially true when models have a dynamic mechanism for selecting planting date, then the fixed growing season will reduce the significance of the relationship between yield and the climate variables. Second, the models do not examine the CFE or impacts of pest or disease under future climate conditions, as it is still difficult to discern this effect from management practice and technological development using historical data. Most of the experiments on impacts of CFE on crop radiation use efficiency are limited to low level of CO_2 concentration (approximately 550ppm), thus available data on CFE at high level of CO_2 are limited. Thirdly, this study could benefit from using more data from other GGCMs and GCMs. The simulated data in this study was limited to the ISIMIP Fast Track, which consists of only 7 GGCMs and 5 GCMs. As ISIMIP/AgMIP project has published results from the third iteration, with more GGCMs and GCM inputs.

In detecting the significant trends of mean temperature, the study may suffer from bias resulting from: i) downscaling methodology for the GCM ensemble, and ii) the choice of model in the ensemble and climate change scenarios. For the former, this study could benefit by employing other downscaling approaches, including other statistical and dynamical downscaling, to reduce the bias caused by the downscaling method. For the latter, as more GCMs are developed and advanced, and as knowledge of climate change scenarios are improved, more data can be used to develop new large-scale observational ensembles.

As the study took advantage of available of data for crop yield and climate to build time-efficient statistical models, other environmental and non-climatic factors that may have an impact on crop yield may be omitted. While the panel linear regression approach with fixed effects allows

the control of time-invariant factors such as soil and county-specific characteristics, non-observed factors may still have a confounding impact on the yield and climate relationship. This can be improved in further studies by: i) collecting more data, and ii) using more complex statistical models such as machine learning methods.

References

Abatzoglou, J. T. 2013. Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology,* 33**,** 121-131.

Abatzoglou, J. T. & Brown, T. J. 2012. A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology,* 32**,** 772-780.

Acharjee, T. K., Halsema, G. V., Ludwig, F. & Hellegers, P. 2017a. Declining trends of water requirements of dry season Boro rice in the north-west Bangladesh. *Agricultural Water Management,* 180**,** 148-159.

Acharjee, T. K., Ludwig, F., Van Halsema, G., Hellegers, P. & Supit, I. 2017b. Future changes in water requirements of Boro rice in the face of climate change in North-West Bangladesh. *Agricultural Water Management,* 194**,** 172-183.

Adachi, S. A., Nishizawa, S., Ando, K., Yamaura, T., Yoshida, R., Yashiro, H., Kajikawa, Y. & Tomita, H. 2018. An evaluation method for uncertainties in regional climate projections. *Atmospheric Science Letters,* 20.

Adnan, S. & Khan, A. H. 2009. Effective rainfall for irrigated agriculture plains of Pakistan. *Pakistan Journal of Meteorology,* 6**,** 61-72.

Ali, M. & Mubarak, S. 2017. Effective Rainfall Calculation Methods for Field Crops: An Overview, Analysis and New Formulation. *Asian Research Journal of Agriculture,* 7**,** 1-12.

Allen Jr, L. 1990. Plant responses to rising carbon dioxide and potential interactions with air pollutants. 19**,** 15-34.

Allen, R. G., Pereira, L. S., Raes, D. & Smith, M. 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome,* 300**,** D05109.

Anapalli, S. S., Green, T. R., Reddy, K. N., Gowda, P. H., Sui, R., Fisher, D. K., Moorhead, J. E. & Marek, G. W. 2018. Application of an energy balance method for estimating evapotranspiration in cropping systems. *Agricultural Water Management,* 204**,** 107-117.

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P. J., Rötter, R. P., Cammarano, D., Brisson, N., Basso, B., Martre, P., Aggarwal, P. K., Angulo, C., Bertuzzi, P., Biernath, C., Challinor, A. J., Doltra, J., Gayler, S., Goldberg, R., Grant, R., Heng, L., Hooker, J., Hunt, L. A., Ingwersen, J., Izaurralde, R. C., Kersebaum, K. C., Müller, C., Naresh Kumar, S., Nendel, C., O'leary, G., Olesen, J. E., Osborne, T. M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M. A., Shcherbak, I., Steduto, P., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J. W., Williams, J. R. & Wolf, J. 2013. Uncertainty in simulating wheat yields under climate change. *Nature Climate Change,* 3**,** 827-832.

Auffhammer, M., Ramanathan, V. & Vincent, J. R. 2011. Climate change, the monsoon, and rice yield in India. *Climatic Change,* 111**,** 411-424.

Ban, N., Schmidli, J. & Schär, C. 2015. Heavy precipitation in a changing climate: Does short-term summer precipitation increase faster? *Geophysical Research Letters,* 42**,** 1165-1172.

Barnes, E. A., Thompson, D. W. J., Deser, C., Foust, W. E. & Phillips, A. S. 2015. Quantifying the Role of Internal Climate Variability in Future Climate Trends. *Journal of Climate,* 28**,** 6443-6456.

Bautista-Capetillo, C., Zavala, M. & Martínez-Cob, A. 2013. Using Thermal Units for Crop Coefficient Estimation and Irrigation Scheduling Improves Yield and Water Productivity of Corn (

Zea mays

L.). *Journal of Irrigation and Drainage Engineering,* 139**,** 214-220.

Blanc, E. 2017a. Aggregation of gridded emulated rainfed crop yield projections at the national or regional level. *Journal of Global Economic Analysis,* 2**,** 112-127.

Blanc, É. 2017b. Statistical emulators of maize, rice, soybean and wheat yields from global gridded crop models. *Agricultural and Forest Meteorology,* 236**,** 145-161.

Blanc, E. & Schlenker, W. 2017. The Use of Panel Models in Assessments of Climate Impacts on Agriculture. *Review of Environmental Economics and Policy,* 11**,** 258-279.

Blanc, E. & Sultan, B. 2015. Emulating maize yields from global gridded crop models using statistical estimates. *Agricultural and Forest Meteorology,* 214-215**,** 134-147.

Bondeau, A., Smith, P. C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-Campen, H., Müller, C., Reichstein, M. & Smith, B. 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology,* 13**,** 679-706.

Brouwer, C. & Heibloem, M. 1986. Irrigation water management: irrigation water needs. *Training manual,* 3.

Brüderl, J. & Ludwig, V. 2015. Fixed-effects panel regression. *The Sage handbook of regression analysis and causal inference***,** 327-357.

Cameron, A. C., Gelbach, J. B. & Miller, D. L. 2011. Robust Inference With Multiway Clustering. *Journal of Business & Economic Statistics,* 29**,** 238-249.

Carrão, H., Russo, S., Sepulcre-Canto, G. & Barbosa, P. 2016. An empirical standardized soil moisture index for agricultural drought assessment from remotely sensed data. *International Journal of Applied Earth Observation and Geoinformation,* 48**,** 74-84.

Challinor, A. J., Ewert, F., Arnold, S., Simelton, E. & Fraser, E. 2009. Crops and climate change: progress, trends, and challenges in simulating impacts and informing adaptation. *J Exp Bot,* 60**,** 2775-89.

Challinor, A. J., Watson, J., Lobell, D. B., Howden, S. M., Smith, D. R. & Chhetri, N. 2014. A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change,* 4**,** 287-291.

Chen, C., Cane, M. A., Wittenberg, A. T. & Chen, D. 2017. ENSO in the CMIP5 Simulations: Life Cycles, Diversity, and Responses to Climate Change. *Journal of Climate,* 30**,** 775-801.

Cheng, Y., Voisin, N., Yearsley, J. R. & Nijssen, B. 2020a. Reservoirs Modify River Thermal Regime Sensitivity to Climate Change: A Case Study in the Southeastern United States. *Water Resources Research,* 56.

Cheng, Y., Voisin, N., Yearsley, J. R. & Nijssen, B. 2020b. Thermal extremes in regulated river systems under climate change: an application to the southeastern U.S. rivers. *Environmental Research Letters,* 15.

Cleveland, W. S. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association,* 74**,** 829-836.

Cleveland, W. S. & Devlin, S. J. 1988. Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association,* 83**,** 596-610.

Climate Action Tracker. 2021. *Temperatures* [Online]. Climate Action Tracker. Available: https://climateactiontracker.org/global/temperatures/ [Accessed 01/19 2022].

Dai, A., Fyfe, J. C., Xie, S.-P. & Dai, X. 2015. Decadal modulation of global surface temperature by internal climate variability. *Nature Climate Change,* 5**,** 555-559.

Dakshinamurti, C., Saxena, P. N. & Rao, E. G. K. 1961. Evaluation of Effective Rainfall and Irrigation from Ground-water Measurements. *Nature,* 190**,** 1100-1101.

De Boeck, H. J., Kimball, B. A., Miglietta, F. & Nijs, I. 2012. Quantification of excess water loss in plant canopies warmed with infrared heating. *Glob Chang Biol,* 18**,** 2860-8.

Department of Economic and Social Affairs, U. N. 2019. *2019 Revision of World Population Prospects* [Online]. Available: https://population.un.org/wpp/ [Accessed 10/23 2020].

Deryng, D., Elliott, J., Folberth, C., Müller, C., Pugh, T. a. M., Boote, K. J., Conway, D., Ruane, A. C., Gerten, D., Jones, J. W., Khabarov, N., Olin, S., Schaphoff, S., Schmid, E., Yang, H. & Rosenzweig, C. 2016. Regional disparities in the beneficial effects of rising CO2 concentrations on crop water productivity. *Nature Climate Change,* 6**,** 786-790.

Deryng, D., Sacks, W. J., Barford, C. C. & Ramankutty, N. 2011. Simulating the effects of climate and agricultural management practices on global crop yield. *Global Biogeochemical Cycles,* 25.

Deser, C., Knutti, R., Solomon, S. & Phillips, A. S. 2012a. Communication of the role of natural variability in future North American climate. *Nature Climate Change,* 2**,** 775-779.

Deser, C., Phillips, A., Bourdette, V. & Teng, H. 2010. Uncertainty in climate change projections: the role of internal variability. *Climate Dynamics,* 38**,** 527-546.

Deser, C., Phillips, A., Bourdette, V. & Teng, H. 2012b. Uncertainty in climate change projections: the role of internal variability. *Climate Dynamics,* 38**,** 527-546.

Deser, C., Terray, L. & Phillips, A. S. 2016. Forced and Internal Components of Winter Air Temperature Trends over North America during the past 50 Years: Mechanisms and Implications*. *Journal of Climate,* 29**,** 2237-2258.

Deser, C. & Wettstein, J. J. 2014. Internal Variability in Projections of Twenty-First-Century Arctic Sea Ice Loss: Role of the Large-Scale Atmospheric Circulation. *Journal of Climate,* 27**,** 527-550.

Djaman, K., Balde, A. B., Sow, A., Muller, B., Irmak, S., N'diaye, M. K., Manneh, B., Moukoumbi, Y. D., Futakuchi, K. & Saito, K. 2015. Evaluation of sixteen reference evapotranspiration methods under sahelian conditions in the Senegal River Valley. *Journal of Hydrology: Regional Studies,* 3**,** 139-159.

Djaman, K., O'neill, M., Owen, C., Smeal, D., Koudahe, K., West, M., Allen, S., Lombard, K. & Irmak, S. 2018. Crop Evapotranspiration, Irrigation Water Requirement and Water Productivity of Maize from Meteorological Data under Semiarid Climate. *Water,* 10.

Drake, B. G., Gonzalez-Meler, M. A. & Long, S. P. 1997. MORE EFFICIENT PLANTS: A Consequence of Rising Atmospheric CO2? *Annu Rev Plant Physiol Plant Mol Biol,* 48**,** 609-639.

Earth System Research Laboratories. 2021. *Trends in Atmospheric Carbon Dioxide* [Online]. Available: https://www.esrl.noaa.gov/gmd/ccgg/trends/ [Accessed 4/24 2021].

Ebrahimpour, M., Rahimi, J., Nikkhah, A. & Bazrafshan, J. 2015. Monitoring Agricultural Drought Using the Standardized Effective Precipitation Index. *Journal of Irrigation and Drainage Engineering,* 141.

Economic Research Service. 2018. *Cash receipts by state, commodity ranking and share of U.S. total, 2018 Nominal (current dollars)* [Online]. https://data.ers.usda.gov/reports.aspx?ID=17843#Pf4e60a9205c8455bb85d465d9c518722_9_17iT0R0x33: United States Department of Agriculture Available: https://data.ers.usda.gov/reports.aspx?ID=17843#Pf4e60a9205c8455bb85d465d9c518722_9_17iT0R0x33 [Accessed 2018].

Elliott, J., Kelly, D., Best, N., Wilde, M., Glotter, M. & Foster, I. 2013. The parallel system for integrating impact models and sectors (pSIMS). 1.

Field, C. B., Barros, V. R., Mach, K. J., Mastrandrea, M. D., Van Aalst, R., Adger, W. N., Arent, D. J., Barnett, J., Betts, R. A. & Bilir, T. E. 2014. Technical summary. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the IPCC*.

Frieler, K., Levermann, A., Elliott, J., Heinke, J., Arneth, A., Bierkens, M. F. P., Ciais, P., Clark, D. B., Deryng, D., Döll, P., Falloon, P., Fekete, B., Folberth, C., Friend, A. D., Gellhorn, C., Gosling, S. N., Haddeland, I., Khabarov, N., Lomas, M., Masaki, Y., Nishina, K., Neumann, K., Oki, T., Pavlick, R., Ruane, A. C., Schmid, E., Schmitz, C., Stacke, T., Stehfest, E., Tang, Q., Wisser, D., Huber, V., Piontek, F., Warszawski, L., Schewe, J., Lotze-Campen, H. & Schellnhuber, H. J. 2015. A framework for the cross-sectoral integration of multi-model impact projections: land use decisions under climate impacts uncertainties. *Earth System Dynamics,* 6**,** 447-460.

Goldblum, D. 2013. Sensitivity of Corn and Soybean Yield in Illinois to Air Temperature and Precipitation: The Potential Impact of Future Climate Change. *Physical Geography,* 30**,** 27-42.

Golian, S., Javadian, M. & Behrangi, A. 2019. On the use of satellite, gauge, and reanalysis precipitation products for drought studies. *Environmental Research Letters,* 14.

Gornall, J., Betts, R., Burke, E., Clark, R., Camp, J., Willett, K. & Wiltshire, A. 2010. Implications of climate change for agricultural productivity in the early twenty-first century. *Philos Trans R Soc Lond B Biol Sci,* 365**,** 2973-89.

Gray, S. B., Dermody, O., Klein, S. P., Locke, A. M., Mcgrath, J. M., Paul, R. E., Rosenthal, D. M., Ruiz-Vera, U. M., Siebers, M. H., Strellner, R., Ainsworth, E. A., Bernacchi, C. J.,

Long, S. P., Ort, D. R. & Leakey, A. D. 2016. Intensifying drought eliminates the expected benefits of elevated carbon dioxide for soybean. *Nat Plants,* 2**,** 16132.

Guo, R., Deser, C., Terray, L. & Lehner, F. 2019. Human influence on winter precipitation trends (1921–2015) over North America and Eurasia revealed by dynamical adjustment. *Geophysical Research Letters,* 46**,** 3426-3434.

Hao, Z. & Singh, V. P. 2015. Drought characterization from a multivariate perspective: A review. *Journal of Hydrology,* 527**,** 668-678.

Hasanuzzaman, M., Nahar, K., Alam, M., Roychowdhury, R. & Fujita, M. 2013. Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. *International journal of molecular sciences,* 14**,** 9643-9684.

Hatfield, J. L., Boote, K. J., Kimball, B. A., Ziska, L. H., Izaurralde, R. C., Ort, D., Thomson, A. M. & Wolfe, D. 2011. Climate Impacts on Agriculture: Implications for Crop Production. *Agronomy Journal,* 103**,** 351-370.

Hawkins, E. & Sutton, R. 2009. The Potential to Narrow Uncertainty in Regional Climate Predictions. *Bulletin of the American Meteorological Society,* 90**,** 1095-1108.

Hawkins, E. & Sutton, R. 2010. The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics,* 37**,** 407-418.

Hawkins, E. & Sutton, R. 2012. Time of emergence of climate signals. *Geophysical Research Letters,* 39**,** n/a-n/a.

Heim Jr, R. R. 2002. A review of twentieth-century drought indices used in the United States. *Bulletin of the American Meteorological Society,* 83**,** 1149-1166.

Idso, K. E., Idso, S. B. J. A. & Meteorology, F. 1994. Plant responses to atmospheric CO2 enrichment in the face of environmental constraints: a review of the past 10 years' research. 69**,** 153-203.

Ihsan, M. Z., El-Nakhlawy, F. S., Ismail, S. M., Fahad, S. & Daur, I. 2016. Wheat Phenological Development and Growth Studies As Affected by Drought and Late Season High Temperature Stress under Arid Environment. *Front Plant Sci,* 7**,** 795.

Iizumi, T., Shiogama, H., Imada, Y., Hanasaki, N., Takikawa, H. & Nishimori, M. 2018. Crop production losses associated with anthropogenic climate change for 1981-2010 compared with preindustrial levels. *International Journal of Climatology,* 38**,** 5405-5417.

Jin, Z., Zhuang, Q., Wang, J., Archontoulis, S. V., Zobel, Z. & Kotamarthi, V. R. 2017. The combined and separate impacts of climate extremes on the current and future US rainfed maize and soybean production under elevated CO2. *Glob Chang Biol,* 23**,** 2687-2704.

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., Wilkens, P. W., Singh, U., Gijsman, A. J. & Ritchie, J. T. 2003. The DSSAT cropping system model. *European Journal of Agronomy,* 18**,** 235-265.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J. F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L. & Vertenstein, M. 2015. The Community Earth System Model (CESM) Large Ensemble Project: A

Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. *Bulletin of the American Meteorological Society,* 96**,** 1333-1349.

Keating, B. A., Carberry, P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., Huth, N. I., Hargreaves, J. N. G., Meinke, H., Hochman, Z., Mclean, G., Verburg, K., Snow, V., Dimes, J. P., Silburn, M., Wang, E., Brown, S., Bristow, K. L., Asseng, S., Chapman, S., Mccown, R. L., Freebairn, D. M. & Smith, C. J. 2003. An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy,* 18**,** 267-288.

Khalil, A. 2013. Effect of climate change on evapotranspiration in Egypt. *Researcher,* 5**,** 7-12.

Knutson, T. R. & Zeng, F. 2018. Model assessment of observed precipitation trends over land regions: Detectable human influences and possible low bias in model trends. *Journal of Climate,* 31**,** 4617-4637.

Kogan, F. N. 1995. Droughts of the late 1980s in the United States as derived from NOAA polar-orbiting satellite data. *Bulletin of the American Meteorological Society,* 76**,** 655-668.

Kucharik, C. J. 2008. Contribution of Planting Date Trends to Increased Maize Yields in the Central United States. *Agronomy Journal,* 100**,** 328-336.

Laepple, T. & Huybers, P. 2014. Ocean surface temperature variability: large model-data differences at decadal and longer periods. *Proc Natl Acad Sci U S A,* 111**,** 16682-7.

Leakey, A. D., Ainsworth, E. A., Bernacchi, C. J., Rogers, A., Long, S. P. & Ort, D. R. 2009. Elevated CO2 effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. *J Exp Bot,* 60**,** 2859-76.

Leakey, A. D., Bishop, K. A. & Ainsworth, E. A. 2012. A multi-biome gap in understanding of crop and ecosystem responses to elevated CO2. *Curr Opin Plant Biol,* 15**,** 228-36.

Lehner, F., Deser, C. & Terray, L. 2017. Toward a new estimate of "time of emergence" of anthropogenic warming: Insights from dynamical adjustment and a large initial-condition model ensemble. *Journal of Climate,* 30**,** 7739-7756.

Leng, G. & Hall, J. 2019. Crop yield sensitivity of global major agricultural countries to droughts and the projected changes in the future. *Sci Total Environ,* 654**,** 811-821.

Lesk, C., Rowhani, P. & Ramankutty, N. 2016. Influence of extreme weather disasters on global crop production. *Nature,* 529**,** 84-7.

Lewis, C. S. & Allen, L. N. 2017. Potential crop evapotranspiration and surface evaporation estimates via a gridded weather forcing dataset. *Journal of Hydrology,* 546**,** 450-463.

Li, Y., Guan, K., Schnitkey, G. D., Delucia, E. & Peng, B. 2019. Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. *Glob Chang Biol,* 25**,** 2325-2337.

Liebig, M. A., Franzluebbers, A. J. & Follett, R. F. J. M. a. G. G. C. a. R. T. G. T. a. O. C. C. a. P., San Diego, Ca 2012. Agriculture and climate change: Mitigation opportunities and adaptation imperatives. 3-11.

Livneh, B., Bohn, T. J., Pierce, D. W., Munoz-Arriola, F., Nijssen, B., Vose, R., Cayan, D. R. & Brekke, L. 2015. A spatially comprehensive, hydrometeorological data set for Mexico, the US, and Southern Canada 1950–2013. *Scientific data,* 2**,** 1-12.

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., Maurer, E. P. & Lettenmaier, D. P. 2013. A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. *Journal of Climate,* 26**,** 9384-9392.

Lobell, D., Cahill, K. & Field, C. 2005. Weather-based forecasts of California crop yields. Lawrence Livermore National Laboratory (LLNL), Livermore, CA.

Lobell, D. B. & Asner, G. P. 2003. Climate and management contributions to recent trends in US agricultural yields. *Science,* 299**,** 1032-1032.

Lobell, D. B. & Asseng, S. 2017. Comparing estimates of climate change impacts from process-based and statistical crop models. *Environmental Research Letters,* 12**,** 015001.

Lobell, D. B., Bänziger, M., Magorokosho, C. & Vivek, B. 2011a. Nonlinear heat effects on African maize as evidenced by historical yield trials. *Nature Climate Change,* 1**,** 42-45.

Lobell, D. B. & Bonfils, C. 2008. The Effect of Irrigation on Regional Temperatures: A Spatial and Temporal Analysis of Trends in California, 1934–2002. *Journal of Climate,* 21**,** 2063-2071.

Lobell, D. B. & Burke, M. B. 2008. Why are agricultural impacts of climate change so uncertain? The importance of temperature relative to precipitation. *Environmental Research Letters,* 3.

Lobell, D. B. & Burke, M. B. 2010. On the use of statistical models to predict crop yield responses to climate change. *Agricultural and Forest Meteorology,* 150**,** 1443-1452.

Lobell, D. B., Cassman, K. G. & Field, C. B. 2009. Crop yield gaps: their importance, magnitudes, and causes. *Annual review of environment and resources,* 34**,** 179-204.

Lobell, D. B. & Field, C. B. 2007. Global scale climate–crop yield relationships and the impacts of recent warming. *Environmental Research Letters,* 2.

Lobell, D. B., Field, C. B., Cahill, K. N. & Bonfils, C. 2006. Impacts of future climate change on California perennial crop yields: Model projections with climate and crop uncertainties. *Agricultural and Forest Meteorology,* 141**,** 208-218.

Lobell, D. B., Hammer, G. L., Mclean, G., Messina, C., Roberts, M. J. & Schlenker, W. 2013. The critical role of extreme heat for maize production in the United States. *Nature Climate Change,* 3**,** 497-501.

Lobell, D. B., Schlenker, W. & Costa-Roberts, J. 2011b. Climate trends and global crop production since 1980. *Science,* 333**,** 616-20.

Lorenzo-Lacruz, J., Vicente-Serrano, S. M., López-Moreno, J. I., Beguería, S., García-Ruiz, J. M. & Cuadrat, J. M. 2010. The impact of droughts and water management on various hydrological systems in the headwaters of the Tagus River (central Spain). *Journal of Hydrology,* 386**,** 13-26.

Lu, J., Carbone, G. J. & Gao, P. 2017. Detrending crop yield data for spatial visualization of drought impacts in the United States, 1895–2014. *Agricultural and Forest Meteorology,* 237-238**,** 196-208.

Martin, D. & Gilley, J. 1993. Irrigation Water Requirements, National Engineering Handbook, Chapter 2, Part 623, 302 pp.

Martínez-Cob, A. 2007. Use of thermal units to estimate corn crop coefficients under semiarid climatic conditions. *Irrigation Science,* 26**,** 335-345.

Martins, I. P., De Faria, R. T., Palaretti, L. F., Dos Santos, M. G. & Fischer Filho, J. A. 2019a. Evapotranspiration and crop coefficient of basil determined by weighing lysimeters. *Horticultura Brasileira,* 37**,** 373-378.

Martins, M. A., Tomasella, J. & Dias, C. G. 2019b. Maize yield under a changing climate in the Brazilian Northeast: Impacts and adaptation. *Agricultural Water Management,* 216**,** 339-350.

Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J. W., Rotter, R. P., Boote, K. J., Ruane, A. C., Thorburn, P. J., Cammarano, D., Hatfield, J. L., Rosenzweig, C., Aggarwal, P. K., Angulo, C., Basso, B., Bertuzzi, P., Biernath, C., Brisson, N., Challinor, A. J., Doltra, J., Gayler, S., Goldberg, R., Grant, R. F., Heng, L., Hooker, J., Hunt, L. A., Ingwersen, J., Izaurralde, R. C., Kersebaum, K. C., Muller, C., Kumar, S. N., Nendel, C., O'leary, G., Olesen, J. E., Osborne, T. M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M. A., Shcherbak, I., Steduto, P., Stockle, C. O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., White, J. W. & Wolf, J. 2015. Multimodel ensembles of wheat growth: many models are better than one. *Glob Chang Biol,* 21**,** 911-25.

Matiu, M., Ankerst, D. P. & Menzel, A. 2017. Interactions between temperature and drought in global and regional crop yield variability during 1961-2014. *PLoS One,* 12**,** e0178339.

Mckinnon, K. A. & Deser, C. 2018. Internal Variability and Regional Climate Trends in an Observational Large Ensemble. *Journal of Climate,* 31**,** 6783-6802.

Mckinnon, K. A., Poppick, A., Dunn-Sigouin, E. & Deser, C. 2017. An "Observational Large Ensemble" to Compare Observed and Modeled Temperature Trend Uncertainty due to Internal Variability. *Journal of Climate,* 30**,** 7585-7598.

Melillo, J. M., Richmond, T. & Yohe, G. J. T. N. C. A. 2014. Climate change impacts in the United States. 52.

Mistry, M. N., Sue Wing, I. & De Cian, E. 2017. Simulated vs. empirical weather responsiveness of crop yields: US evidence and implications for the agricultural impacts of climate change. *Environmental Research Letters,* 12**,** 075007.

Moore, F. C. & Lobell, D. B. 2015. The fingerprint of climate trends on European crop yields. *Proc Natl Acad Sci U S A,* 112**,** 2670-5.

Morison, J. I. L. 1985. Sensitivity of stomata and water use efficiency to high CO2. *Plant, Cell and Environment,* 8**,** 467-474.

Muchow, R. C., Sinclair, T. R. & Bennett, J. M. 1990. Temperature and Solar Radiation Effects on Potential Maize Yield across Locations. *Agronomy Journal,* 82**,** 338.

Murphy, L. N., Bellomo, K., Cane, M. & Clement, A. 2017. The role of historical forcings in simulating the observed Atlantic multidecadal oscillation. *Geophysical Research Letters,* 44**,** 2472-2480.

Nelsen, R. B. 2007. *An introduction to copulas*, Springer Science & Business Media.

Nicholls, N. 1997. Increased Australian wheat yield due to recent climate trends. *Nature,* 387**,** 484-485.

Noaa. 2020. *Billion-Dollar Weather and Climate Disasters: Summary Stats* [Online]. Available: https://www.ncdc.noaa.gov/billions/summary-stats/US/1980-2019 [Accessed September 07 2020].

Norby, R. J. & Zak, D. R. 2011. Ecological Lessons from Free-Air CO2 Enrichment (FACE) Experiments. *Annual Review of Ecology, Evolution, and Systematics,* 42**,** 181-203.

Okumura, Y. M., Tomas, R. A., Phillips, A. S., Deser, C., Alexander, M. A., Capotondi, A., Scott, J. D., Kwon, Y.-O. & Ohba, M. 2012. ENSO and Pacific Decadal Variability in the Community Climate System Model Version 4. *Journal of Climate,* 25**,** 2622-2651.

Ostberg, S., Schewe, J., Childers, K. & Frieler, K. 2018. Changes in crop yields and their variability at different levels of global warming. *Earth System Dynamics,* 9**,** 479-496.

Oyebamiji, O. K., Edwards, N. R., Holden, P. B., Garthwaite, P. H., Schaphoff, S. & Gerten, D. 2015. Emulating global climate change impacts on crop yields. *Statistical Modelling,* 15**,** 499-525.

Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., Church, J. A., Clarke, L., Dahe, Q. & Dasgupta, P. 2014. *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*, Ipcc.

Painter, J. A., Torak, L. J. & Jones, J. W. 2015. Evaluation and comparison of methods to estimate irrigation withdrawal for the National Water Census Focus Area Study of the Apalachicola-Chattahoochee-Flint River Basin in southwestern Georgia.

Peña-Gallardo, M., Vicente-Serrano, S. M., Quiring, S., Svoboda, M., Hannaford, J., Tomas-Burguera, M., Martín-Hernández, N., Domínguez-Castro, F. & El Kenawy, A. 2019. Response of crop yield to different time-scales of drought in the United States: Spatio-temporal patterns and climatic and environmental drivers. *Agricultural and Forest Meteorology,* 264**,** 40-55.

Pereira, L. S., Allen, R. G., Smith, M. & Raes, D. 2015. Crop evapotranspiration estimation with FAO56: Past and future. *Agricultural Water Management,* 147**,** 4-20.

Pervez, M. S. & Brown, J. F. 2010. Mapping Irrigated Lands at 250-m Scale by Merging MODIS Data and National Agricultural Statistics. *Remote Sensing,* 2**,** 2388-2412.

Peters, A. J., Rundquist, D. C. & Wilhite, D. A. 1991. Satellite detection of the geographic core of the 1988 Nebraska drought. *Agricultural and Forest Meteorology,* 57**,** 35-47.

Petersen, L. 2019. Impact of Climate Change on Twenty-First Century Crop Yields in the U.S. *Climate,* 7.

Phillips, A. S., Deser, C. & Fasullo, J. 2014. Evaluating Modes of Variability in Climate Models. *Eos, Transactions American Geophysical Union,* 95**,** 453-455.

Pielke, R. A., Pitman, A., Niyogi, D., Mahmood, R., Mcalpine, C., Hossain, F., Goldewijk, K. K., Nair, U., Betts, R., Fall, S., Reichstein, M., Kabat, P. & De Noblet, N. 2011. Land use/land cover changes and climate: modeling analysis and observational evidence. *WIREs Climate Change,* 2**,** 828-850.

Quiring, S. M. & Papakryiakou, T. N. 2003. An evaluation of agricultural drought indices for the Canadian prairies. *Agricultural and Forest Meteorology,* 118**,** 49-62.

Ralston Fonseca, F., Jaramillo, P., Bergés, M. & Severnini, E. 2019. Seasonal effects of climate change on intra-day electricity demand patterns. *Climatic Change,* 154**,** 435-451.

Ramírez, J. A., Finnerty, B. J. J. O. I. & Engineering, D. 1996. CO 2 and temperature effects on evapotranspiration and irrigated agriculture. 122**,** 155-163.

Rayner, N. A. 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research,* 108.

Rhee, J., Im, J. & Carbone, G. J. 2010. Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data. *Remote Sensing of Environment,* 114**,** 2875-2887.

Roberts, M. J., Braun, N. O., Sinclair, T. R., Lobell, D. B. & Schlenker, W. 2017. Comparing and combining process-based crop models and statistical models with some implications for climate change. *Environmental Research Letters,* 12**,** 095010.

Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S. & Mosher, S. 2013. Berkeley Earth Temperature Averaging Process. *Geoinformatics & Geostatistics: An Overview,* 01.

Rosenzweig, C., Arnell, N. W., Ebi, K. L., Lotze-Campen, H., Raes, F., Rapley, C., Smith, M. S., Cramer, W., Frieler, K., Reyer, C. P. O., Schewe, J., Van Vuuren, D. & Warszawski, L. 2017. Assessing inter-sectoral climate change risks: the role of ISIMIP. *Environmental Research Letters,* 12.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A. C., Müller, C., Arneth, A., Boote, K. J., Folberth, C., Glotter, M. & Khabarov, N. 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the National Academy of Sciences,* 111**,** 3268-3273.

Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P., Antle, J. M., Nelson, G. C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G. & Winter, J. M. 2013. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. *Agricultural and Forest Meteorology,* 170**,** 166-182.

Ruane, A. C., Rosenzweig, C., Asseng, S., Boote, K. J., Elliott, J., Ewert, F., Jones, J. W., Martre, P., Mcdermid, S. P., Muller, C., Snyder, A. & Thorburn, P. J. 2017. An AgMIP framework for improved agricultural representation in IAMs. *Environ Res Lett,* 12.

Saadi, S., Todorovic, M., Tanasijevic, L., Pereira, L. S., Pizzigalli, C. & Lionello, P. 2015. Climate change and Mediterranean agriculture: Impacts on winter wheat and tomato crop evapotranspiration, irrigation requirements and yield. *Agricultural Water Management,* 147**,** 103-115.

Sacks, W. J., Deryng, D., Foley, J. A., Ramankutty, N. J. G. E. & Biogeography 2010. Crop planting dates: an analysis of global patterns. 19**,** 607-620.

Sanchez-Lorenzo, A., Morán-Tejeda, E., Revuelto, J., Azorin-Molina, C., López-Moreno, J. I., Camarero, J. J., Lorenzo-Lacruz, J., Beguería, S. & Vicente-Serrano, S. M. 2012. Performance of Drought Indices for Ecological, Agricultural, and Hydrological Applications. *Earth Interactions,* 16**,** 1-27.

Sánchez, N., González-Zamora, Á., Piles, M. & Martínez-Fernández, J. 2016. A New Soil Moisture Agricultural Drought Index (SMADI) Integrating MODIS and SMOS Products: A Case of Study over the Iberian Peninsula. *Remote Sensing,* 8.

Santer, B. D., Bonfils, C. J. W., Fu, Q., Fyfe, J. C., Hegerl, G. C., Mears, C., Painter, J. F., Po-Chedley, S., Wentz, F. J., Zelinka, M. D. & Zou, C.-Z. 2019. Celebrating the anniversary of three key events in climate change science. *Nature Climate Change,* 9**,** 180-182.

Scarpare, F. V., Rajagopalan, K., Liu, M., Nelson, R. L. & Stöckle, C. O. 2022. Evapotranspiration of Irrigated Crops under Warming and Elevated Atmospheric CO2: What Is the Direction of Change? *Atmosphere,* 13.

Schauberger, B., Archontoulis, S., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Elliott, J., Folberth, C., Khabarov, N., Muller, C., Pugh, T. A., Rolinski, S., Schaphoff, S., Schmid, E., Wang, X., Schlenker, W. & Frieler, K. 2017. Consistent negative response of US crops to high temperatures in observations and crop models. *Nat Commun,* 8**,** 13931.

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., Dankers, R., Eisner, S., Fekete, B. M., Colon-Gonzalez, F. J., Gosling, S. N., Kim, H., Liu, X., Masaki, Y., Portmann, F. T., Satoh, Y., Stacke, T., Tang, Q., Wada, Y., Wisser, D., Albrecht, T., Frieler, K., Piontek, F., Warszawski, L. & Kabat, P. 2014. Multimodel assessment of water scarcity under climate change. *Proc Natl Acad Sci U S A,* 111**,** 3245-50.

Schlenker, W. & Lobell, D. B. 2010. Robust negative impacts of climate change on African agriculture. *Environmental Research Letters,* 5**,** 014010.

Schlenker, W. & Roberts, M. J. 2009. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proc Natl Acad Sci U S A,* 106**,** 15594-8.

Shi, W., Tao, F. & Zhang, Z. 2013. A review on statistical models for identifying climate contributions to crop yields. *Journal of Geographical Sciences,* 23**,** 567-576.

Shindell, D. & Faluvegi, G. 2009. Climate response to regional radiative forcing during the twentieth century. *Nature Geoscience,* 2**,** 294-300.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K. & Venevsky, S. 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology,* 9**,** 161-185.

Smith, C. A., Scott, J. D., Phillips, A. S., Vimont, D. J., Schneider, N., Nakamura, H., Minobe, S., Miller, A. J., Mantua, N. J., Di Lorenzo, E., Deser, C., Cobb, K. M., Ault, T. R., Alexander, M. A. & Newman, M. 2016. The Pacific Decadal Oscillation, Revisited. *Journal of Climate,* 29**,** 4399-4427.

Smith, M. 1992. *CROPWAT: A computer program for irrigation planning and management*, Food & Agriculture Org.

Spiliotopoulos & Loukas 2019. Hybrid Methodology for the Estimation of Crop Coefficients Based on Satellite Imagery and Ground-Based Measurements. *Water*, 11.

Stocker, T. 2014. *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change*, Cambridge University Press.

Stöckle, C. O., Donatelli, M. & Nelson, R. 2003. CropSyst, a cropping systems simulation model. *European journal of agronomy,* 18**,** 289-307.

Stone, P. 2001. The effects of heat stress on cereal yield and quality. *Crop responses and adaptations to temperature stress***,** 243-291.

Stott, P. A., Gillett, N. P., Hegerl, G. C., Karoly, D. J., Stone, D. A., Zhang, X. & Zwiers, F. 2010. Detection and attribution of climate change: a regional perspective. *WIREs Climate Change,* 1**,** 192-211.

Sunil, A., Deepthi, B., Mirajkar, A. B. & Adarsh, S. 2020. Modeling future irrigation water demands in the context of climate change: a case study of Jayakwadi command area, India. *Modeling Earth Systems and Environment,* 7**,** 1963-1977.

Tack, J., Barkley, A. & Nalley, L. L. 2015. Effect of warming temperatures on US wheat yields. *Proc Natl Acad Sci U S A,* 112**,** 6931-6.

Tao, F., Yokozawa, M. & Zhang, Z. 2009. Modelling the impacts of weather and climate variability on crop productivity over a large area: A new process-based model development, optimization, and uncertainties analysis. *Agricultural and Forest Meteorology,* 149**,** 831-850.

Tao, X.-E., Chen, H., Xu, C.-Y., Hou, Y.-K. & Jie, M.-X. 2015. Analysis and prediction of reference evapotranspiration with climate change in Xiangjiang River Basin, China. *Water Science and Engineering,* 8**,** 273-281.

Tasumi, M., Allen, R. G. & Trezza, R. Calibrating satellite-based vegetation indices to estimate evapotranspiration and crop coefficients. In Proceedings of the 2006 USCID Water Management Conference, Ground water and surface water under stress: competition, interaction, solutions, 2006. 103-112.

Taub, D. 2010. Effects of rising atmospheric concentrations of carbon dioxide on plants. *Nature Education Knowledge,* 1.

Taylor, K. E., Stouffer, R. J. & Meehl, G. A. 2012. An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society,* 93**,** 485-498.

Tebaldi, C. & Knutti, R. 2007. The use of the multi-model ensemble in probabilistic climate projections. *Philos Trans A Math Phys Eng Sci,* 365**,** 2053-75.

Tebaldi, C. & Lobell, D. B. 2008. Towards probabilistic projections of climate change impacts on global crop yields. *Geophysical Research Letters,* 35.

Tian, L., Yuan, S. & Quiring, S. M. 2018. Evaluation of six indices for monitoring agricultural drought in the south-central United States. *Agricultural and Forest Meteorology,* 249**,** 107-119.

Tigkas, D., Vangelis, H. & Tsakiris, G. 2018. Drought characterisation based on an agriculture-oriented standardised precipitation index. *Theoretical and Applied Climatology,* 135**,** 1435-1447.

Tilman, D., Balzer, C., Hill, J. & Befort, B. L. 2011. Global food demand and the sustainable intensification of agriculture. *Proc Natl Acad Sci U S A,* 108**,** 20260-4.

Trnka, M., Hlavinka, P., Semerádová, D., Dubrovsky, M., Zalud, Z. & Mozny, M. 2007. Agricultural drought and spring barley yields in the Czech Republic. *Plant Soil and Environment,* 53**,** 306.

Tuteja, N. & Singh, G. S. 2012. *Plant acclimation to environmental stress,* Springer Science & Business Media.

Urban, D. W., Sheffield, J. & Lobell, D. B. 2015. The impacts of future climate and carbon dioxide changes on the average and variability of US maize yields under two emission scenarios. *Environmental Research Letters,* 10.

Vanuytrecht, E., Raes, D., Willems, P. & Geerts, S. 2012. Quantifying field-scale effects of elevated carbon dioxide concentration on crops. *Climate Research,* 54**,** 35-47.

Vicente-Serrano, S. M., Beguería, S. & López-Moreno, J. I. 2010. A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate,* 23**,** 1696-1718.

Vittoz, P., Cherix, D., Gonseth, Y., Lubini, V., Maggini, R., Zbinden, N. & Zumbach, S. 2013. Climate change impacts on biodiversity in Switzerland: A review. *Journal for Nature Conservation,* 21**,** 154-162.

Vorosmarty, C. J., Mcintyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R. & Davies, P. M. 2010. Global threats to human water security and river biodiversity. *Nature,* 467**,** 555-61.

Wahid, A., Gelani, S., Ashraf, M. & Foolad, M. 2007. Heat tolerance in plants: An overview. *Environmental and Experimental Botany,* 61**,** 199-223.

Wane, S. & Nagdeve, M. 2014. Estimation of evapotranspiration and effective rainfall using CROPWAT. *International Journal of Agricultural Engineering,* 7**,** 23-26.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O. & Schewe, J. 2014. The Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP): project framework. *Proc Natl Acad Sci U S A,* 111**,** 3228-32.

Wheeler, T. & Von Braun, J. 2013. Climate change impacts on global food security. *Science,* 341**,** 508-13.

White, G. F. & Haas, J. E. 1975. Assessment of research on natural hazards.

Wilcox, J. & Makowski, D. 2014. A meta-analysis of the predicted effects of climate change on wheat yields using simulation studies. *Field Crops Research,* 156**,** 180-190.

Wilhite, D. A. 2000. Chapter 1 Drought as a Natural Hazard: Concepts and Defnitions. Drought Mitigation Center Faculty Publications. 69.

Wilhite, D. A. & Glantz, M. H. 1985. Understanding: the drought phenomenon: the role of definitions. *Water international,* 10**,** 111-120.

Wriedt, G., Van Der Velde, M., Aloe, A. & Bouraoui, F. 2009. Estimating irrigation water requirements in Europe. *Journal of Hydrology,* 373**,** 527-544.

Yihdego, Y., Vaheddoost, B. & Al-Weshah, R. A. 2019. Drought indices and indicators revisited. *Arabian Journal of Geosciences,* 12.

Yusuke, K. 2019. *The Economic Impacts of Drought on US Agriculture* [Online]. Resources Magazine. Available: https://www.resources.org/archives/economic-impacts-drought-us-agriculture/ [Accessed 2021].

Zarei, A. R. & Moghimi, M. M. 2019. Modified version for SPEI to evaluate and modeling the agricultural drought severity. *Int J Biometeorol,* 63**,** 911-925.

Zhongming, Z., Linong, L., Wangqiang, Z. & Wei, L. 2019. Climate Change and Land-An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems.

Zipper, S. C., Qiu, J. & Kucharik, C. J. 2016. Drought effects on US maize and soybean production: spatiotemporal patterns and historical changes. *Environmental Research Letters,* 11.

Appendix A Study region historical climate trend

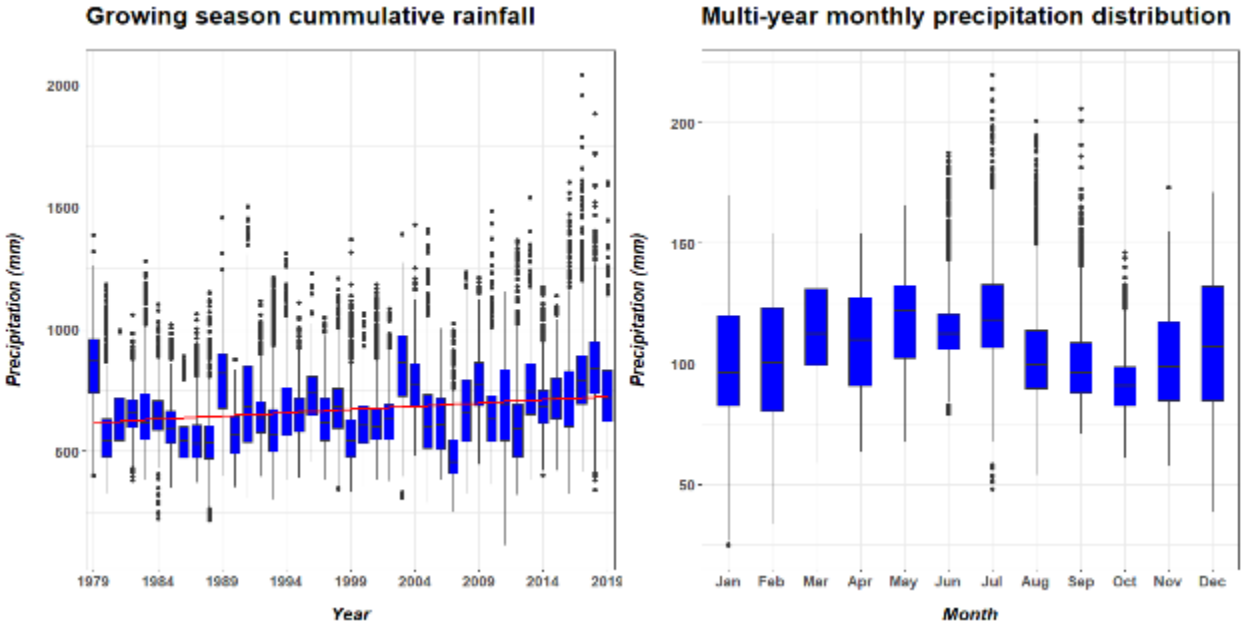

Figure A 1 Time series of growing season data (left) and distribution of monthly average (right) over the 1979-2019 period for cumulative precipitation. The boxplots encompass data from all counties in the SERC, the black line is the median, the box represents interquartile range, and the extended whisker includes 90% of the data points. The red line is the trend of the time series over the study period.

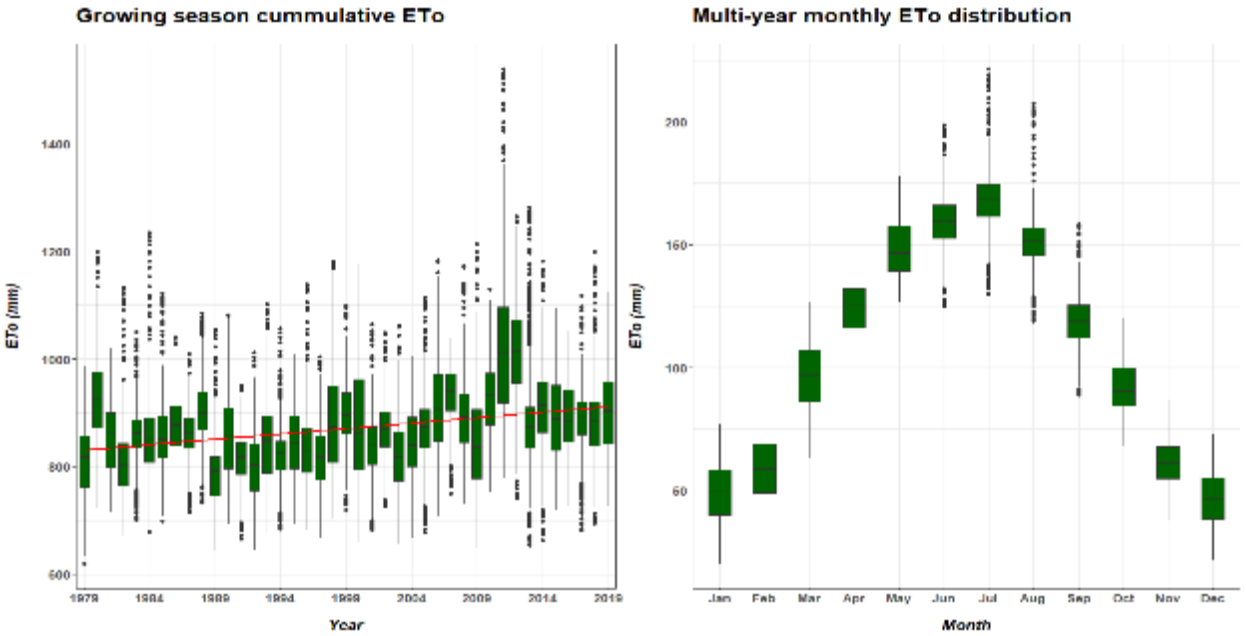

Figure A 2 Same as Figure A 1 but for cumulative reference evapotranspiration (ET_o)

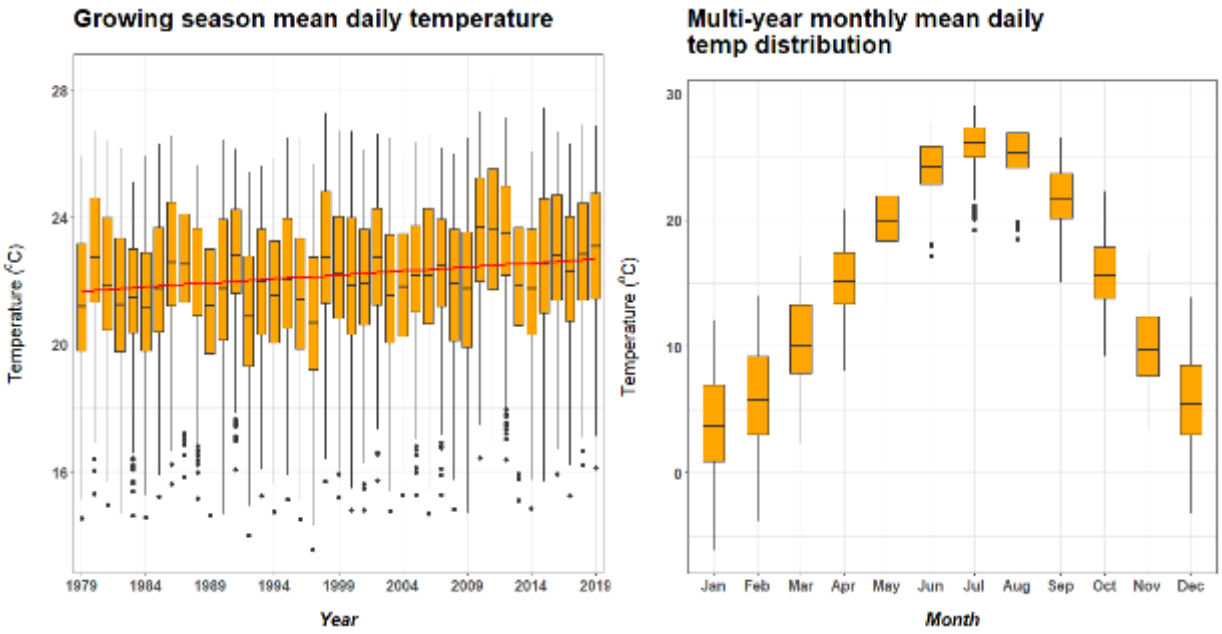

Figure A 3 Same as Figure A 1 but for daily mean temperature

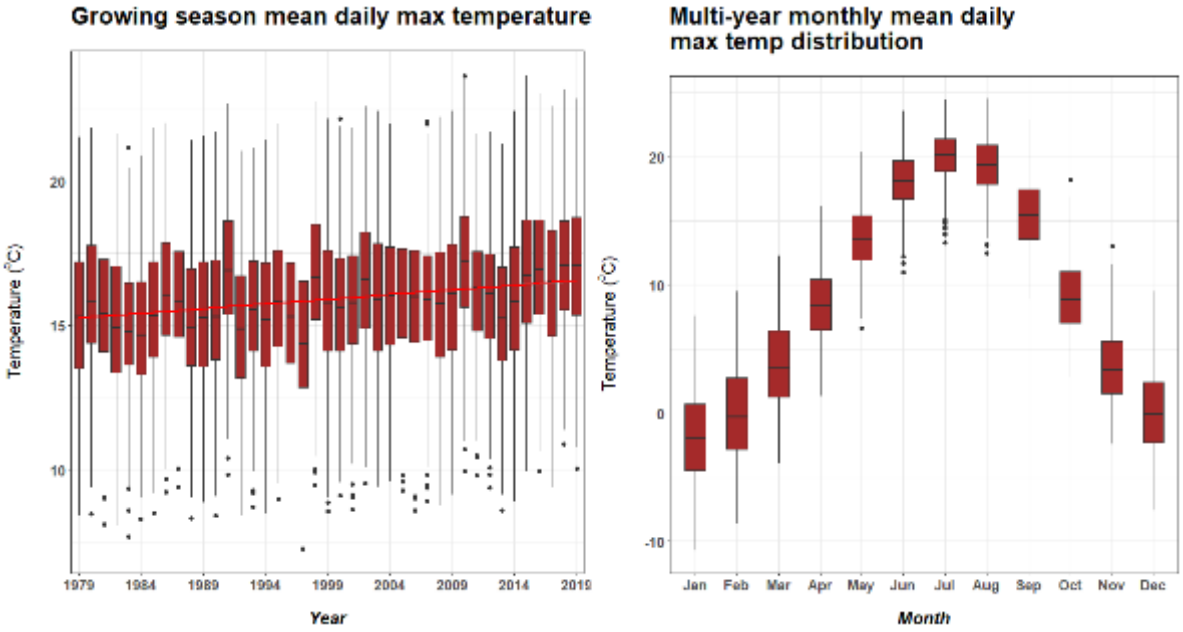

Figure A 4 Same as Figure A 1 but for daily maximum temperature

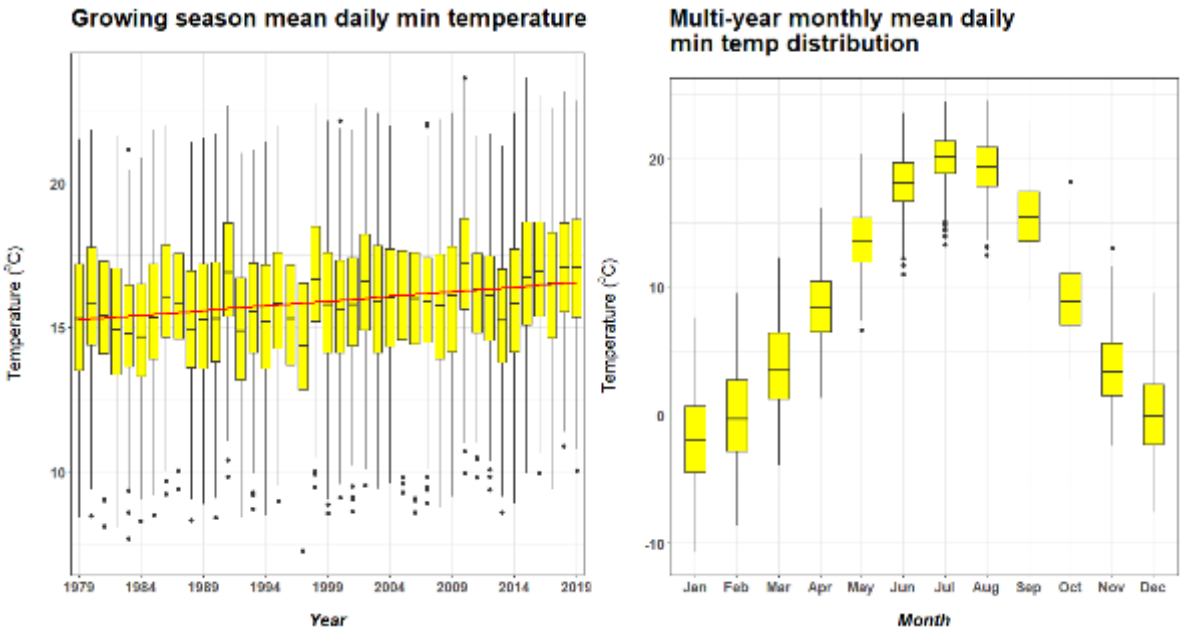

Figure A 5 Same as Figure A 1 but for but for daily minimum temperature

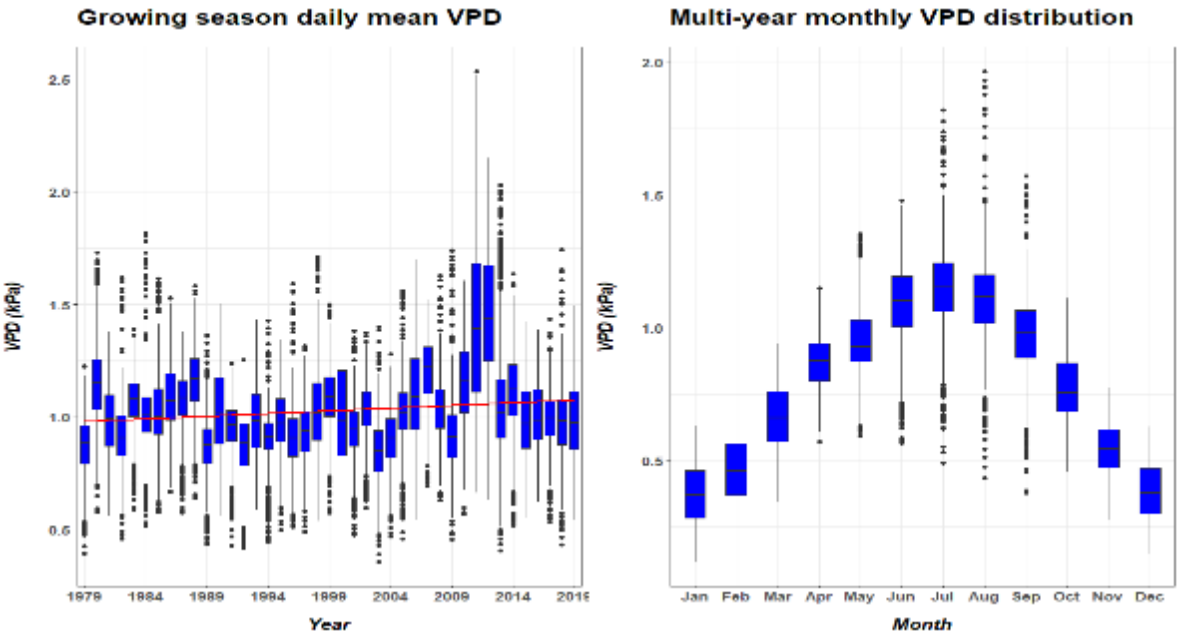

Figure A 6 Same as Figure A 1 but for daily vapor pressure deficit.

Appendix B Global gridded crop models

Table B 1 Global gridded crop models (GGCMs) from the AgMIP projects used for developing the emulators

GGCM	***Version***	***Institution***	***Contact person / Web address***
LPJmL	-	Potsdam Institute for Climate Impact Research	Stefan Olin/Thomas Pugh Stefan.Olin@nateko.lu.se thomas.pugh@imk.fzk.de
pDSSAT	1.0	University of Chicago Computation Institute	Joshua Elliott, jelliott@ci.uchicago.edu
PEGASUS	1.1	Tyndall Centre University of East Anglia, UK / McGill University, Canada	Delphine Deryng d.deryng@uea.ac.uk

Table B 2 Summary of Stress factors in GGCMs

GGCM	**LPJmL**	**pDSSAT**	**PEGASUS**
Scale	Ecosystem	Site-based	Ecosystem
Spatial scale	0.5º lon x 0.5º lat		
Temporal scale	Daily simulation, yearly yield		
Climate input	Tmean, P, percentage of cloud cover (or Radiation)	Tmax, Tmin P, Radiation	Tmean, Tmax, Tmin, P, percentage of cloud cover
Soil input	Harmonized world soil database; Soil texture classification based on the USDA soil texture classification; thermal parameters	Harmonized world soil database	Available Water Capacity
Factor of Stresses	Water, Temperature	Water, Temperature, Specific-heat stress, Oxygen stress, Nitrogen stress	Water, Temperature, Specific-heat stress, Nitrogen stress, Phosphorus stress, Potassium stress
Type of heat stress	N/A	Vegetative; Reproductive organ; Number of grain (pod) set during the flowering period	Reproductive organ; Number of grain (pod) set during the flowering period
Type of water stress	Soil available water in root zone	Ratio of supply to demand of water	Ratio of supply to demand of water
Fertilizer application	N/A	Dynamic application	Annual application

Table B 3 Summary of Biophysical processes in GGCMs

GGCM	LPJmL	pDSSAT	PEGASUS
Leaf area development	Prescribed shape of LAI curve as function of phenology, modified by water stress & low productivity	Dynamic simulation based on development and growth processes	Dynamic simulation based on development and growth processes
Yield formation	Harest Index modified by water stress	Number of grains and grain growth rate	Partitioning during reproductive stages
Crop phenology is a function of	Temperature; Vernalization	Temperature; Vernalization; Photoperiod; Other water/nutrient stress effects	Temperature (Heat unit index)
Evapo-transpiration	Priestley –Taylor	Priestley –Taylor	Priestley –Taylor
Soil water dynamic (number of soil layers in the model)	5	4	3
Soil CN model	N/A	Carbon, Nitrogen, Phosphorous model (3 organic matter pools)	N/A
CO2 effect	Leaf-level photosynthesis; Stomatal conductance	Radiation use efficiency (for soybean: Leaf-level photosynthesis); Transpiration efficiency	Radiation use efficiency; Transpiration efficiency
CO2 baseline	370 ppm (2000)	330 ppm (1975)	369 ppm (2000)
Planting date	Fixed planting day after 1951	Fixed planting window	Dynamic planting window (adaptation to climate change)
Crop cultivars	GDD* requirements and vernalization requirements computed based on past climate experience	Simulate crop GDDs requirement according to estimated annual GDDs from daily temperature	Dynamic GDD requirement (adaptation to climate change)

*GDD: growing degree days

Table B 4 Summary of yield change projections under climate change scenarios by the models for maize (Unit of yield change is percentage)

Model	GCM	2030-2059			2070-2099		
		Min	Median	Max	Min	Median	Max
Observation	GFDL	-30.59	-20.63	-7.27	-60.52	-46.83	-21.2
	HadGEM2	-56.69	-44.97	-18.39	-79.66	-72.04	-48.25
	NorESM1	-37	-26.13	-8.87	-61.62	-50.98	-28.52
LPJmL_m	GFDL	-13.72	-5.81	-0.06	-35.47	-24.87	-10.47
	HadGEM2	-34.53	-20.99	1.8	-45.72	-37.34	-24.46
	NorESM1	-18	-5.53	4.4	-28.81	-15.83	-4.95
pDSSAT_m	GFDL	-18.89	-7.66	-0.45	-36.79	-27.87	-16.79
	HadGEM2	-34.48	-25.73	-9.8	-50.38	-43.66	-32.11
	NorESM1	-22.06	-11.71	-1.91	-34.82	-25.23	-16.79
PEGASUS_m	GFDL	-26.31	-17.17	-12.06	-49.74	-40.29	-28.36
	HadGEM2	-46.22	-38.71	-25.23	-67.95	-62.37	-48.45
	NorESM1	-31.62	-23.37	-15.11	-52.04	-44.35	-35.63

Table B 5 Summary of yield change projections under climate change scenarios by the models for soybeans (Unit of yield change is percentage)

Model	GCM	2030-2059			2070-2099		
		Min	Median	Max	Min	Median	Max
Observation	GFDL	-23.05	-10.12	4.23	-46.98	-31.83	-0.19
	HadGEM2	-42.82	-30.61	1.96	-66.82	-55.32	-16.88
	NorESM1	-26.94	-14.9	6.82	-45.58	-33.48	-0.6
LPJmL_m	GFDL	-9.63	-2.75	3.96	-23.78	-16.98	-6.7
	HadGEM2	-23.36	-14.61	0.4	-31.21	-25.03	-14.32
	NorESM1	-13.76	-4.13	3.97	-19.52	-10.56	-1.83
pDSSAT_m	GFDL	-61.69	-43.65	-29.33	-91.92	-82.06	-66.34
	HadGEM2	-89.82	-78.19	-41.74	-98.09	-96	-89.82
	NorESM1	-67.63	-49.12	-29.16	-90.72	-80.8	-69.87
PEGASUS_m	GFDL	-33.66	-27.21	-23.6	-61.51	-54.8	-47.65
	HadGEM2	-59.3	-52.54	-37.29	-79.06	-75.87	-61.73
	NorESM1	-43.11	-34.72	-25.65	-66.64	-59.37	-49.78

VITA

PERSONAL INFORMATION	
Name:	Hai (Heidi) Nguyen
Current Location:	Raleigh, NC

EDUCATION		
School	**Degree (Level & Concentration)**	**Years**
Pennsylvania State University	**PhD of Agricultural and Biological Engineering**	2016-2022
University of Leuven (KU Leuven), Leuven, Belgium	**Master of Science in Water Resources Engineering**	2012-2014
Thuy Loi University (TLU), Hanoi, Vietnam	**Bachelor of Hydrological and Environmental Engineering**	2004-2009

Professional Experience:

USDA-ARS, State College, PA **July 2020 – January 2022**
Agricultural Data Scientist Internship

- Designed and developed a machine learning model to discover the pattern of and predict soil nitrogen emission from growing bioenergy crops on marginal land based on field experimental data with better accuracy level than previous studies.
- Designed, processed, and discovered spatial pattern of pesticide use and environmental impacts in the U.S. crop production though USDA-NASS survey data, providing suggestions to farmers to reduce the use of harmful pesticides.
- Analyzed and visualized impacts of field management of miscanthus and switchgrass on marginal land on nitrous oxide emission using R.

Pennsylvania State University, State College, PA **January 2020 – January 2022**
Graduate Research Associate - Department of Agricultural and Biological Engineering

- Assessed drought risks of food crops over the Southeastern region of the U.S using big data analysis techniques.
- Examined impacts on climate change on crop growth, including potential for earlier planting, shortened growing season and changes in dry-wet conditions due to warmer climate.
- Predicted future crop water demand for food crops in the Southeastern region of the U.S under climate change impacts.
- Detected emergent trends in ecologically-relevant climate variables to understand biodiversity exposure to changing environments through innovative and robust large ensemble approach.

University of Missouri, Columbia, MO **June 2016 – December 2019**
Graduate Research Associate
Department of Biological, Bioenvironmental & Chemical Engineering

- Designed and executed statistical models to estimate climate change impacts on crop productivity and water demand as part of a $4-million NSF-funded interdisciplinary project.
- Collected, managed, and processed big data of climate, including data from a large ensemble of global circulation models (GCMs) in an efficient manner, using R, Python, ArcGIS-Arcpy and high-performance computing system.
- Built an ensemble of statistical models using large climate and yield datasets to assess impacts of climate change on the crop production over the southeastern US, using statistical and spatial analysis
- Discovered an innovative and efficient modelling approach to utilize the available big data from global agricultural modeling projects.