Advancing the Modeling of Climate Impacts on Annual and Perennial Crops to Inform Adaptation Pathways: Lessons from California Almonds and U.S. Maize

By

SHUAIQI WU

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Atmospheric Science

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Erwan Monier

Paul Ullrich

Yufang Jin

Committee in Charge

2025

Contents

Abstract

Agriculture faces intensifying challenges from climate change, threatening both perennial and annual crops that are crucial to global food and nutrient security. This dissertation advances the modeling of climate impacts on agriculture by developing and applying a novel statistical framework for perennial crops, and by scaling and calibrating a process-based model for annual crops, thereby improving yield projections under climate change and informing adaptation planning.

The first study develops a novel modeling framework that integrates climate modeling, horticultural science, and statistical yield modeling, and applies it to California almonds. Results show that increasing minimum temperatures and humidity during the bloom and pollination stage, along with heat stress during the growing stage, are primary drivers of yield losses. Climate change is projected to reduce almond yields by up to 49% by 2100 under the high warming scenario (SSP585). However, sustained innovation gains could more than offset climate damages, highlighting the joint role of technological progress and climate adaptation in perennial systems.

The second study expands AquaCrop, a process-based crop model developed by the Food and Agriculture Organization, for regional-scale applications. A flexible Python-based gridded implementation is developed and calibrated using two decades of county-level U.S. maize yield data, which substantially improves model performance compared to the default field-scale calibrated AquaCrop. This county-level calibration approach enables AquaCrop to capture both spatial and temporal dynamics of observation data more effectively than existing approaches, demonstrating the value of regional calibration for large-scale applications and establishing a scalable framework for future continental and global assessments.

The third study applies the county-level calibrated AquaCrop to examine U.S. maize production under climate change. Driven by CMIP6 climate projections, simulations indicate that U.S. maize production could decline by 17% by the end of the century under the high warming scenario (SSP585), with particularly severe impacts in the Corn Belt region. Analyses of adaptation strategies show that irrigation, fertilization, and spatial relocation of maize systems can potentially offset climate damages, but feasibility is constrained by water availability, land competition, and resource efficiency. County-level projections provide actionable insights for agricultural policy, crop insurance, and farm-level planning.

Together, these studies advance crop modeling for both annual and perennial systems, demonstrating new approaches to evaluate climate risks and identify potential adaptation strategies. By bridging methodological innovation with practical case studies of California almonds and U.S. maize, the dissertation provides insights that are broadly applicable to assessing climate impacts and informing adaptation strategies across diverse crops and regions.

Acknowledgments

First of all, I would like to express my sincere gratitude to my advisor, Prof. Erwan Monier, for his support, guidance, and encouragement throughout my Ph.D journey. His advices are invaluable not only for my academic growth but also my personal development. I could not have asked for a better advisor.

Next, I'd like to thank my committee members for their thoughtful feedback and generous support throughout the process. I am also grateful to all my collaborators whose contributions significantly strengthened my work. In addition, I extend my thanks to the faculty members and students in the atmospheric science graduate group. I'm especially thankful to my friend, Dr. Shiheng Duan, who has always been willing to share his insights and experiences, which have greatly broadened my prospective and enriched my study.

Last but not least, I 'd like to express my deepest gratitude to my parents, my wife, and my sister. I couldn't achieve this accomplishment without their love, support, and sacrifice. This achievement belongs to them as much as it does to me.

Introduction

Food and nutrient insecurity is one of the greatest global challenges, with more than 2 billion people experiencing moderate to severe food insecurity in 2023 (FAO, 2024). Pressures from population growth, climate change, and dwindling water resources intensify this crisis (Molotoks et al., 2021; Singh et al., 2017; Wheeler & Von Braun, 2013; Zhao et al., 2017). Evidence shows that rising temperatures, shifting precipitation, and more frequent extremes are already depressing yields and increasing volatility across major crops, with elevated risks of simultaneous production shocks in multiple breadbaskets (Iizumi et al., 2018; Lesk et al., 2016; Lobell et al., 2011; Ray et al., 2015; Tigchelaar et al., 2018). Yet most assessments emphasize staple annual crops such as wheat, maize, soybeans, and rice (Haqiqi et al., 2021; Lobell et al., 2020; Lobell & Gourdji, 2012; Rosenzweig et al., 2014; Wing et al., 2015), while impacts on perennial crops, such as nuts and fruits that provide vital nutrition and substantial economic benefits, remain comparatively understudied (Kreitzman et al., 2020; Leisner, 2020; Olba-Zięty et al., 2021; Shi et al., 2023). Therefore, understanding climate impacts on both annual and perennial crops is essential to guide effective adaptation and safeguard food and nutrition security.

In this context, crop models serve as pivotal tools to support research, farm-level decisions, and policymaking. Statistical crop models find relationships between yield and climate variables and use these relationships to predict yield under different climate conditions (Lobell et al., 2006). Without the strict limitation of data inputs, a wide range of variables can be fed to the model. In addition to weather data and their derivatives, soil characteristics, and water availability, which are present in most data-driven analysis, cross-disciplinary variables can be included as well to investigate the interaction between agriculture and other fields. Moreover, statistical modeling is more widely used to simulate the influence of climate on the yield of perennial crops because

it is challenging for process-based models to account for long lifespans, multi-year reproductive cycles, and complex physiology of perennial crops (Zhu et al., 2021). However, statistical crop models typically lack interpretability and transparency because they do not explicitly represent crop physiological processes, which limits their utility for understanding biological mechanisms (C. Kumar et al., 2025; Paudel et al., 2023). In addition, most studies using statistical models to simulate corp yield do not identify the actual crop physiological processes involved in the climate effects. Therefore, new approaches are needed to support adaptation to climate change for crops

Process-based crop models simulate the growth of crops with a daily to hourly time step and generate a series of crop variables, such as yield and biomass, based on equations and theories from crop physiology, phenology, soil science, and biometeorology (Roberts et al., 2017). Although the structure of process-based crop models can vary, they usually incorporate multiple modules, which include but are not limited to plant development, soil-water balance, weather, and farming management, to comprehensively represent the underlying physical process of crop growth and the interaction of crop, environment, and farming system (Jones et al., 2003; Keating et al., 2003; Steduto et al., 2009).Despite their extensive application in crop simulations, process-based crop models face multiple challenges (Pasquel et al., 2022). For example, inconsistent calibration procedures are a major source of calibration uncertainty due to variations in parameter selection, evaluation metrics, and calibration algorithms (Pasquel et al., 2022; Wallach et al., 2021). The lack of extensive, complete, and consistent observational data is another major limitation, especially for regional calibration (Grassini et al., 2015). The majority of process-based crop models were originally developed as site-based to accurately simulate the crop growth process at the field scale and their initial parameterizations were mostly conducted at the field or plot scale (Seidel et al., 2018; Wallach et al., 2021). However, parameters obtained from site-specific calibration usually fail to account for the spatial heterogeneity in crop growth environment, cultivar characteristics, and management practices, which casts doubts on models' accuracy and robustness for large-scale application (Ramirez-Villegas et al., 2017). On the other hand, process-based models are often run at coarse spatial resolutions, which limit their ability to inform stakeholders. As a result, there is a need to develop flexible and transparent modeling frameworks that stakeholders can use at higher resolutions but across large scales.

To assess climate change impacts on agriculture and explore adaptation options, we develop two

approaches. First, we build a statistical modeling framework for perennial crops by linking crop yields with key optimal and damaging environmental conditions associated with various growth stages, and apply it to California almonds, which supply about 80% of the world's almonds and provide robust data (California Department of Food and Agriculture, 2022). Second, we scale the stakeholder-focused process-based model AquaCrop to the continental scale and calibrate it using U.S. maize as a case study, because maize is the most extensively cultivated and consumed cereal globally, and U.S. leads in maize production by volume and economic value (FAO, 2024; USDA NASS, 2025). Together, these efforts advance crop modeling of climate impacts in annual and perennial systems and provide critical insights for adaptation.

Chapter 1

Advancing the Modeling of Future Climate and Innovation Impacts on Perennial Crops to Support Adaptation: A Case Study of California Almonds

Wu, S., Zikalala, P. G., Alba, S., Jarvis-Shean, K. S., Kisekka, I., Segaran, M., Snyder, R., Monier, E. (2025). Advancing the modeling of future climate and innovation impacts on perennial crops to support adaptation: A case study of California almonds. *Earth's Future*, 13, e2024EF005033. https://doi.org/10.1029/2024EF005033

Abstract

Perennial crops are vital to the global food supply, providing valuable nutrition and economic benefits, but are at risk of severe climate damages. Most climate research has focused on major annual crops like cereals and has focused on the overall impact of climate change on yields providing limited actionable knowledge to support adaptation. In this study, we bring together climate scientists, biometeorology specialists, plant scientists, and agricultural engineers to develop a new perennial crop modeling framework that integrates climate modeling, horticulture and agronomy science, and statistical modeling. We apply this framework to California almonds as a case study, because they offer robust data to calibrate and evaluate our model. Our model quantifies the influence of climate in each almond development stage and of innovation on county-level yields. We simulate future yield changes under a large multi-model ensemble of high-resolution climate simulations and innovation scenarios. We find that climate change could lead to yield losses of 17% by 2100 under moderate warming (SSP245) and 49% under high warming (SSP585); however, we also find that sustained innovation gains could more than offset these negative climate impacts. We identify increasing minimum temperatures and humidity during the bloom and pollination period as well as heat stress during the growing period as the main drivers of yield losses. We discuss synergistic strategies to limit the negative impacts of climate change and to ensure continued gains from innovation. This modeling approach could provide valuable insights into climate adaptation strategies for other perennial crops and regions.

1.1 Introduction

Climate change threatens current and future crop production and poses a risk to global food security (Gregory et al., 2005; Tubiello et al., 2007; Wheeler & Von Braun, 2013). Most studies examining the impact of climate change on agriculture focus on staple annual crops like wheat, maize, soybeans, and rice, because these crops are the main source of calories for the global population (Haqiqi et al., 2021; Lobell et al., 2020; Lobell & Gourdji, 2012; Rosenzweig et al., 2014; Wing et al., 2015). In contrast, there is limited research on the impacts of future climate change on perennial crops like nuts and fruits although they serve as a vital source of nutrition and provide significant economic benefits to society (Kreitzman et al., 2020; Leisner, 2020; Olba-Zięty et al., 2021; Shi et al., 2023).

The current knowledge gaps in understanding climate-related damages to perennial crops pose risks to global food and nutrition security and hinder growers' ability to develop and implement effective adaptation strategies. To address this, our study advances the modeling of future climate change and innovation impacts on perennial crops and uses California almonds as a case study. The Mediterranean climate, the soil, and the geographical conditions of the Central Valley of California have historically produced high yields for specialty crops, including almond; however climate change is presenting new challenges that threaten the California almond industry. Ranked second among agricultural goods produced in California by value, almonds provided cash receipts of $ 5.03 billion, approximately 10% of the total revenue generated from Californian agriculture in 2021 (California Department of Food and Agriculture, 2022). Besides accounting for 99% of the domestic supply, the California almond industry produced about 80% of the world's almonds and contributed the highest export value to California among all other agricultural export commodities in 2021. Additionally, more than 1.2 million jobs were supported by the almond industry in California in 2016 (California Department of Food and Agriculture, 2022). The importance of California almonds to the global almond supply and to the agriculture sector of California, coupled with the availability of robust data, drive our decision to select almonds in California as the case study. Better understanding the exact pathways involved in climate damages to perennial crops is crucial to identifying and supporting adaptation strategies to ensure the resilience of the agriculture sector.

Developing process-based models representing the impact of climate on almond yield is challenging because process-based models generally struggle to simulate the periodic defoliation and subsequent regrowth of perennial crops (Maestrini et al., 2022). Therefore, empirical modeling is more widely used to simulate the influence of climate on the yield of perennial crops. Previous research has indicated certain statistical relationships between almond yield and climate; however, most studies only focus on key variables like temperature and precipitation or on seasonal averages, thus not providing direct linkages to almond phenology. For example, previous studies have relied solely on average minimum daily temperature in February and average January precipitation (Lobell et al., 2006) or on seasonal average maximum daily temperature and precipitation (Hong et al., 2020) as predictors of California almond yields. In both studies, little insight is provided on how each variable considered influences the almond phenological development stages. In fact, almond phenological stages do not always match seasons, and yields can be influenced by numerous environmental

factors beyond average temperature and precipitation, like humidity, extreme heat, and wind speed (Brittain et al., 2013; Connell, 2000; Palacio-Bielsa et al., 2017). As a result, current studies provide limited actionable information to help growers and production agriculture researchers understand the major climate drivers that impact almond yield and to support potential adaptation strategies.

In California, the annual life cycle of commonly grown almond varieties encompasses dormancy stage from November to January, pollination and bloom from February to mid-March, nut growth stage from mid-March to June, hull split stage from June to July, and the harvest between August and October (Almond Board of California, n.d.). During the dormancy period, almond trees must experience a certain amount of winter chill to meet the cultivars' specific threshold to emerge from dormancy and bloom in the spring (Benmoussa et al., 2017; Hayhoe et al., 2004; Luedeling et al., 2009; Rodríguez et al., 2019). The success of pollination is crucial to almond production and honeybee colonies play a crucial role in transporting almond pollen from one cultivar to another because of the self-incompatibility of most commercial almond varieties (Connell, 2000; Ortega et al., 2004). Unfavorable environmental conditions such as cold temperatures (below 12.8°C) and high wind speed (higher than 24 km/hr) can hinder the foraging performance of honeybees, compromising almond harvest (Connell, 2000). The almond pollen physiology including anther dehiscence, germination, and tube growth is also sensitive to temperature with specific optimum temperature ranges (Connell, 2000). A strong negative correlation between almond yield and February minimum daily temperature was identified by (Lobell & Field, 2011), of which the underlying mechanism is associated with higher temperatures shortening the pollination window and impacting the stigma's reception to pollen (Lobell & Field, 2011; Ortega et al., 2004). Fungal and bacterial diseases, such as blossom and twig blight, and delayed dehiscence are serious concerns for almond fruit growth especially with high humidity or precipitation during the bloom phase (Gradziel, 2009). After the bloom and before harvest, heat accumulation controls the growth of the almond fruit and nut (Tombesi et al., 2010). Excessive (or insufficient) heat accumulation can lead to advancing (or delaying) the fruit maturation, thus impeding the harvest schedule (Benmoussa et al., 2017; Campoy et al., 2011; Ramírez et al., 2015). Finally, precipitation during the harvest period can delay the timing of harvest and increase the risk of post-harvest microbial contamination (Hayman & Thomas, 2017; Palacio-Bielsa et al., 2017; Thomas, 2019).

In addition to being impacted by climate conditions, historical almond yields in California have

shown steady increases in the past four decades, a direct consequence of continuous efforts to improve farming practices and advance agronomic technology like pest management, fertilizer application, and new rootstock varieties (Buchner et al., 2001; Epstein et al., 2001; Ewert et al., 2005; Muhammad et al., 2015; Romero et al., 2004). Meanwhile, cultivation of drought tolerant almond varieties, regulated deficit irrigation, and evaporative cooling are studied to adapt to negative influences such as water scarcity and extreme heat from climate change (Freitas et al., 2023; Jahanzad et al., 2020; Martínez-García et al., 2022; Parker et al., 2020). For these reasons, it is important to frame potential future impacts of climate change on almond yields in the context of an industry experiencing major gains in yield from innovation, which no studies to date have done for future California almonds. However, because there is limited historical information on the introduction of improved farming practices and new agronomic technologies, explicitly modeling future innovation and its impact on future yields is challenging and most studies rely on innovation scenarios (Alston et al., 1995; Askari & Cummings, 1977; Ewert et al., 2005).

Previous studies using statistical models to simulate crop yield often fail to capture the complex physiological and phenological processes that drive crop development and yield. Meanwhile, future yield gains from innovation are usually overlooked in yield projections for perennial crops. To remedy these shortcomings, we rethink the traditional disciplinary approaches to modeling the impact of climate change on agriculture and bring together climate scientists, biometeorology specialists, plant scientists, and agricultural engineers to integrate climate modeling, horticulture and agronomy science, as well as statistical modeling and machine learning to further our understanding of how the human-Earth systems interactions shape the future of perennial crops and identify possible adaptation strategies (Reed, Hadjimichael, Moss, Brelsford, et al., 2022; Reed, Hadjimichael, Moss, Monier, et al., 2022). We are dealing with the complex integration of human-driven systems (agriculture and innovation) and natural systems (climate changes), which are interacting dynamically in time and space, and thus require novel modeling approaches. Specifically, we develop a Least Absolute Shrinkage and Selection Operator (LASSO) regression modeling framework that links county-level almond yield to 14 different agro-climate indices (ACIs) (Monier et al., 2016) that represent the main optimal and damaging environmental conditions associated with the various almond growth stages and to temporal trends representing yield gains from innovation. The inclusion of phenology-derived ACIs enables the model to quantify the impact of climate change directly on each phenological

development stage and provide actionable information to develop targeted adaptation strategies to limit the negative influence of climate change. We build 1000 LASSO regression models by randomly selecting 70% of the data 1000 times to estimate confidence intervals in the statistical relationship between almond yield and ACIs and test each model on the remaining 30%. We then simulate future almond yields under two high-resolution multi-model ensembles of downscaled simulations with 8 CMIP6 (Eyring et al., 2016) and 18 CMIP5 (Taylor et al., 2012) climate models, and under two sets of climate scenarios correspondingly, namely Shared Socioeconomic Pathways (SSP, (Riahi et al., 2017)) and Representative Concentration Pathways (RCP, (Van Vuuren et al., 2011)). Finally, we attempt to incorporate future innovation in our modeling framework by assuming three innovation scenarios based on historical almond yield trends: a sustained innovation scenario that assumes continued historical yield trends; a decelerating innovation scenario that assumes innovation exponentially plateaus by 2100; a climate change only scenario, where no additional innovation is assuming past 2020.

1.2 Data and methods

1.2.1 County-level Almond Data in California

The United States Department of Agriculture (USDA) National Agricultural Statistics Service archived and published detailed county-level crop statistics, including yields, production, harvested areas, and prices (USDA NASS, 2023). The recorded crop data dates back to 1980, which provides us with more than 40 years of observed crop yields for major agricultural products in California. A total of 16 counties where almond trees are planted are mostly located in the Central Valley region: Tehama, Butte, Glenn, Yuba, Colusa, Sutter, Yolo, Solano, San Joaquin, Stanislaus, Madera, Merced, Fresno, Tulare, Kinds, and Kern.

1.2.2 Climate Data

To examine the historical influence of weather and climate on observed California crop yields, we require high-resolution historical gridded climate information. We use the gridMET dataset for that purpose (Abatzoglou, 2013). GridMET is derived from the PRISM observational climate data (Daly et al., 2008) with desirable temporal attributes (and additional variables) from the NLDAS-2

regional reanalysis (Mitchell et al., 2004), and provides daily gridded surface meteorological variables at a spatial resolution of 1/24th degree (∼4 km) covering the contiguous United States from 1979 to present.

To estimate the impact of climate change on California crop yield, we use two large ensembles of high-resolution future climate projections. We rely on the 3-km resolution version of the Localized Constructed Analogs Version 2 (LOCA2-Hybrid) climate ensemble developed specifically for California and that consists of 15 CMIP6 Global Climate Models (GCMs) downscaled using the Livneh high-resolution historical gridded climate dataset (Livneh et al., 2015) as observational data for training purposes (Pierce et al., 2023, 2014). LOCA2-Hybrid provides both historical (1950-2014) and future (2015-2100) climate projections under SSP245, SSP370, and SSP585 scenarios. We also use the Multivariate Adaptive Constructed Analogs (MACA) version 2 climate ensemble that consists of 20 CMIP5 GCMs downscaled at 4-km resolution using GridMET as the baseline observational dataset (Abatzoglou & Brown, 2012), also commonly referred to as MACAv2-METDATA. MACAv2-METDATA projects both historical (1950-2005) and future (2006-2099) climate variables under RCP 4.5 and RCP 8.5 scenarios. Because some downscaled GCM simulations lack the variables needed in this study, we use 8 CMIP6 GCMs from LOCA2-Hybrid and 18 CMIP5 GCMs from MACAv2-METDATA. Although assumptions used to develop the SSP and RCP scenarios are different, SSP245 and SSP585 have similar forcing levels as RCP4.5 and RCP8.5 respectively and they are representative medium and high warming scenarios in their own pathways(O'Neill et al., 2016).

1.2.3 Almond Cropland Layer

In previous studies developing statistical models linking climate to almond yield at the county level, climate information is averaged over the entire county, including regions where crops have not been grown in the historical record, like foothills and mountains. Here, we use the geospatial Cropland Data Layer (CDL), developed by the USDA, a raster dataset containing historical geographical locations of croplands across the contiguous U.S. at a 30-m resolution (Boryan et al., 2011). We use the CDL dataset to identify grid cells where almonds are grown and to calculate the county-level average climate that almond orchards have been exposed to eliminate regions where almonds have not been grown historically. Since the CDL dataset for California is only available after 2007, we

assume that almond orchards prior to 2007 were generally located within the boundaries of the planting areas of the past 14 years (2007-2020).

1.2.4 ACI calculation

Leveraging the expertise in crop physiology, horticulture, and agronomy of our interdisciplinary team, we identify the optimal and damaging environmental conditions in each phenological growth stage, which we refer to as agro-climate indices (ACI). These ACIs can be adapted to other perennial crops and potentially to annual crops by carefully tailoring them to the specific phenology and physiology of the crop considered. It is worth mentioning that ACIs may vary across large spatial scales so ACIs may need to be modified when applied to different regions of the world. The ACIs used in this study (Table 1) are summarized as follows:

Table 1.1: Agro-Climatic Index list

Dormancy	November to January	1. Dormancy_Chill (Benmoussa et al., 2017; Palacio-Bielsa et al., 2017) 2. Dormancy_ETo (Spinelli et al., 2018; Xue et al., 2021) 3. Jan_P (Lobell et al., 2007)
Bloom (pollination)	February to March (First Half)	1. Bloom_P (Connell, 2000; Palacio-Bielsa et al., 2017) 2. Bloom_Tmin (Lobell et al., 2007; Ortega et al., 2004) 3. Bloom_FrostDays (Traynor, 2017) 4. Bloom_ETo (Spinelli et al., 2018; Xue et al., 2021) 5. Bloom_GDD (Connell, 2000) 6. Bloom_Humidity (Palacio-Bielsa et al., 2017) 7. Bloom_WindyDays (Brittain et al., 2013; Connell, 2000)
Growing	March to June	1. Growing_GDD(Parker & Abatzoglou, 2018; Rattigan & Hill, 1986) 2. Growing_KDD(Freitas et al., 2023; Parker et al., 2020) 3. Growing_ETo(Spinelli et al., 2018; Xue et al., 2021)
Hull split	June to July	
Harvest	August to October	1. Harvest_P(Palacio-Bielsa et al., 2017; Thomas, 2019)

Dormancy_Chill: The cumulative chill degree hours during dormancy are calculated by applying

the trigonometric concept and following the method of (Baldocchi & Wong, 2008):

$$ChillDegreeHours = 2 * \frac{6}{T_{ave} - T_{min}} * (T_{ref} - T_{min}) \tag{1.1}$$

ETo: Given the available meteorological variables provided by the LOCA2-Hybrid and MACAv2-METDATA climate datasets, the averaged Reference Evapotranspiration (ETo) is calculated through the FAO Penman-Monteith equation for dormancy, bloom, and growing periods (R. G. Allen et al., 1998). Daily ETo data are available in the gridMET dataset but we calculate daily ETo for both LOCA2-Hybrid and MACAv2-METDATA.

Bloom_FrostDays: Sum of days with daily minimum temperature below 0 °C during the bloom period.

P: Accumulated precipitation (P) over dormancy, bloom, and harvest periods.

Bloom_Tmin: Averaged daily minimum temperature of bloom period.

Bloom_GDD: The cumulative growing degree days (GGD) with base temperature of 4.5 °C and upper temperature of 30 °C for the bloom period are calculated based on the single sine method (J. C. Allen, 1976).

Bloom_Humidity: Averaged daily humidity over bloom period.

Bloom_WindyDays: Sum of days with wind speed greater than 24 km/hr.

Growing_GDD: Cumulative growing degree days (GDD) are calculated using the method as Bloom_GDD with a base temperature of 4.5 °C and a upper temperature of 35 °C for the growing period.

Growing_KDD: Cumulative killing degree days (KDD) are computed using the method as Bloom_GDD but with a base temperature of 35 °C.

The ACIs chosen are largely uncorrelated (Fig. S1).

1.2.5 Statistical Modeling

We select the LASSO regression method to build the statistical relationship between yearly county-level almond yields and ACIs because LASSO can successfully mitigate the potential multicollinearity (Chong & Jun, 2005) and provide direct relationships between yields and each ACI. Multicollinearity is a concept that refers to when several predictor variables in a model are highly correlated ($R^2 > 0.7$),

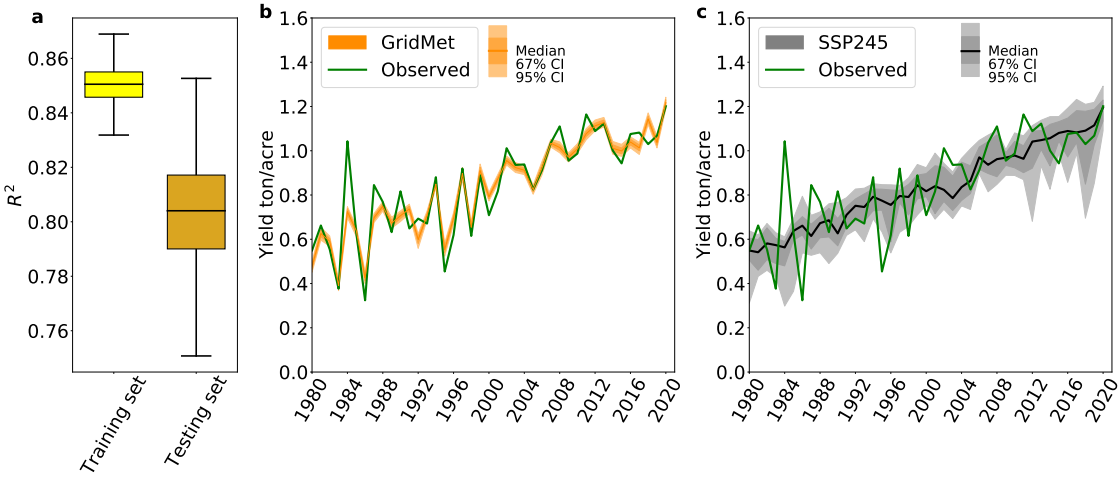

Figure 1.1: Evaluation of statistical models. Coefficients of determination (R2) obtained from training dataset (70%) and testing dataset (30%) for the 1000 statistical models are shown in **a**. **b** and **c** display the California average yield time series simulated by the 1000 statistical models driven by GridMet climate and by LOCA2-Hybrid climate over the historical period (1980-2020). Dark (light) shadings represent the 67% (95%) confidence interval.

which can result in less reliable statistical inferences. LASSO applies the regularization technique to assign a penalty to the summation of magnitudes of the coefficients, thereby minimizing the multicollinearity. We teased out outliers and extreme values in data inputs (ACIs) based on Cook's distance and standardized them by removing the mean and dividing them by their standard deviation (Cook, 2000). We trained a total of 1000 statistical models over the 1980-2020 period by randomly selecting a subset of 70% of the observational dataset and evaluating on the remaining 30%. The Lasso regression model is shown below:

$$Y_c(t) = F_c + \gamma_c t + \sum_i^N [\alpha ACI_c^i(t) + \beta ACI_c^i(t)^2] + \varepsilon_c(t) \tag{1.2}$$

where subscripts c, i, and t indicate county, ACI, and year, respectively; $Y_c(t)$ is the almond yield for county c and year t; F_c refers to the county-level fixed effects representing average yield differences between counties and thus, for example, overall differences in soil quality between counties; γ_c represents county-level temporal trends that cannot be explained by climate variations and generally indicate innovation in agriculture; α and β are the linear and quadratic coefficients; ε_c is the error term.

An evaluation of the statistical models over the historical period driven by the gridMET observed climate (Fig. 1, panel a) shows high coefficients of determination (R2) for both the training set

(median R2=0.85) and the testing set (median R2=0.80) and the ability to reproduce the trends and the year-to-year fluctuations in almond yields, thus demonstrating the robust performance of the statistical modeling approach. Our modeling approach yields higher R2 than the most recent study that simulates county-level almond yields (Hong et al., 2020) who achieve a training R2 close to 0.78 and a testing R2 close to 0.58, which provides additional evidence that our approach is effective. We note that when driven by the LOCA2-Hybrid simulated climate, in some years (i.e. 1984), observed yields fall outside the simulated range, reflecting the well-known inability of global coupled climate models to capture the exact year-to-year chronological climate fluctuations observed in the real world. However, the simulated yields accurately reproduce the observed trend and statistics of year-to-year fluctuations.

1.2.6 Innovation scenarios

We consider three innovation scenarios corresponding to yield increases associated with agronomic innovation such as improved pest management, deficit irrigation, precision agriculture, introduction of new varieties, etc. The sustained innovation scenario assumes that advancement of innovation will continue at the same pace as historical gains (1980-2020), of which the model includes innovation variables that increase by 1 unit per year and range from 1 at 1980 to 120 at 2099. The decelerating innovation scenario assumes the increment of innovation variable decreases by year, reaching a plateau by 2099 with the following equation:

$$Variable_{year} = Variable_{year-1} + \frac{(2099 - year)}{2099 - 2020} \tag{1.3}$$

for *year* ranges from 2021 to 2099 (future period). The stopped innovation scenario assumes no further agronomic innovation after 2020, of which innovation variable increases from 1 at 1980 to 41 for 2020 and stays at 41 for the future years.

1.2.7 Yield change analysis

When calculating impacts of climate change, we are considering a constant innovation level fixed at the year of 2020 to capture only the impact of climate change and to remove any influences from changes in yield associated with innovation trend.

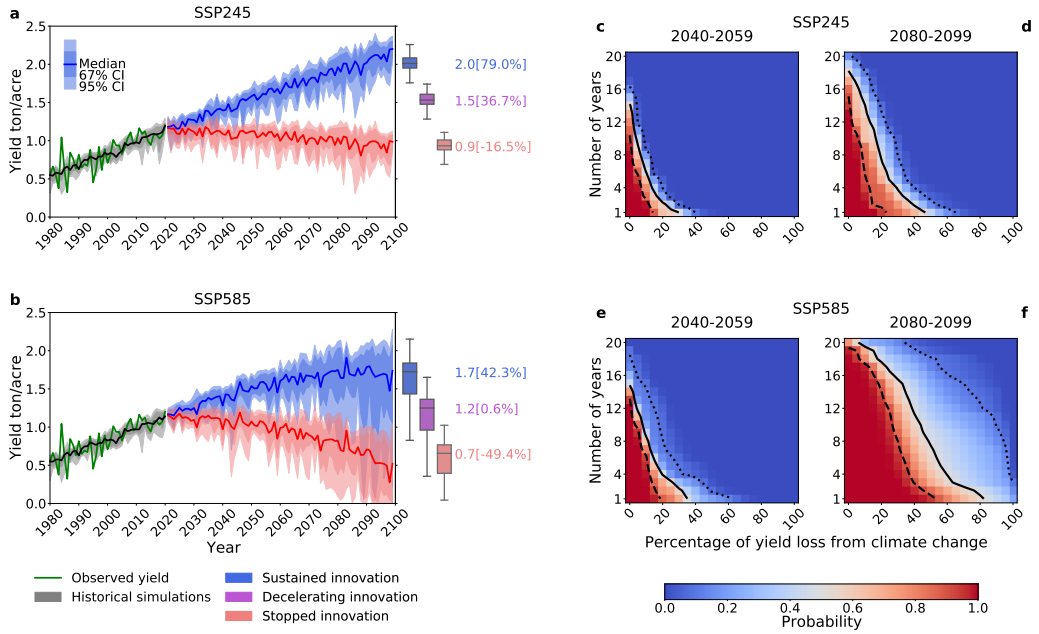

Figure 1.2: Left panel: projections of historical (1980-2020) and future (2021-2099) California state-wide yields under the SSP245 (**a**) and the SSP585 (**b**) LOCA2-Hybrid climate scenarios for three innovation scenarios: stopped innovation (**red**), decelerating innovation (**purple**), and sustained innovation (**blue**). Box plots show distributions of end-of-century (2080-2099) yield predictions, followed by medians of yields and yield changes (shown in bracket) compared with the historical period (2001-2020). Dark (light) shadings represent the 67% (95%) confidence interval from the simulations with 8 GCM6 climate models and 1000 statistical models. Right panel: joint probabilistic distributions of different levels of yield losses and number of years at each loss level for SSP245 scenario (**c** and **d**) and SSP585 scenario (**e** and **f**) for periods of 2040-2059 and 2080-2099. The dotted, solid, and dashed lines represent curves of probabilities of 10%, 50%, and 90%

1.3 Results

1.3.1 Future state-wide yield projections

Climate change is projected to negatively impact California almond yields by 2100 if growers do not adopt effective adaptation strategies, but sustained innovation would offset the damages from climate change (Fig. 2). Under SSP585, climate change alone is projected to decrease yields by approximately 49% for the median simulation by 2100 compared to the historical period (2001-2020) but assuming the same innovation rates as in the past four decades, these yield losses would be

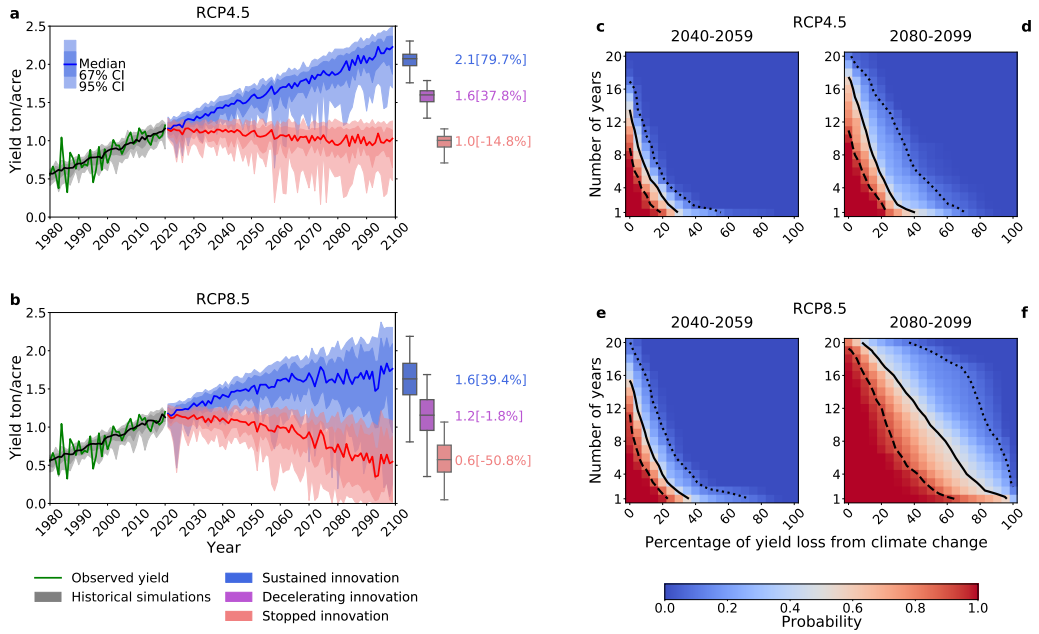

Figure 1.3: Left panel: projections of historical (1980-2020) and future (2021-2099) California state-wide yields under the RCP4.5 (**a**) and the RCP8.5 (**b**) MACAv2-METDATA climate scenarios for three innovation scenarios: stopped innovation (**red**), decelerating innovation (**purple**), and sustained innovation (**blue**). Box plots show distributions of end-of-century (2080-2099) yield predictions, followed by medians of yields and yield changes (shown in bracket) compared with the historical period (2001-2020). Dark (light) shadings represent the 67% (95%) confidence interval from the simulations with 18 GCM5 climate models and 1000 statistical models. Right panel: joint probabilistic distributions of different levels of yield losses and number of years at each loss level for RCP4.5 scenario (**c** and **d**) and RCP8.5 scenario (**e** and **f**) for periods of 2040-2059 and 2080-2099. The dotted, solid, and dashed lines represent curves of probabilities of 10%, 50%, and 90%

more than offset with a net increase by about 42%. Even under the decelerating innovation scenario, yield losses would be completely offset for the median simulation. Under a milder warming scenario (SSP245), climate change alone would lead to yield decreases by 17% while both sustained and decelerating innovation scenarios would far offset these losses with overall gains between 37 and 79% by 2100 for the median simulation. Our analysis demonstrates the large uncertainty in future changes in Californian almond yields associated with different levels of future innovation scenarios, and to a lesser degree with future warming levels. In addition, there are significant variations in yield changes across simulations with climate projections from different climate models and with different

LASSO regression models. For example, while the median yield loss from climate change is 49% under SSP585, the 95th confidence interval ranges from 22 to 90%. Under the decelerating innovation scenario approximately 3000 simulations out of 8,000 show that innovation is not enough to offset the impact of climate change. We explore these sources of uncertainty in detail in subsequent sections. Under climate change alone, a number of simulations projects near total losses in specific years by the end of the century under SSP585, thus indicating severe climate risks to the California almond industry under a high warming scenario. Finally, without future climate change, innovation could increase yields by 155% by 2100 assuming the same innovation rates as in the past four decades, and 113% under the decelerating innovation scenarios.

The yield simulations using the MACAv2-METDATA ensemble under the RCP4.5 and RCP 8.5 scenarios (Fig. 3) provide nearly identical results as with the LOCA2-Hybrid ensemble under the SSP245 and SSP585 scenarios. Since the two climate ensembles vary in the climate models considered, the downscaling technique used, the underlying observational meteorological datasets used for the downscaling, and the scenarios, the consistency in yield projections indicates that the projected yield changes are robust and that the statistical modeling framework is appropriate.

1.3.2 Volatility of projected almond yield

Climate change is also projected to substantially increase year-to-year yield fluctuations, especially in the second half of the 21st century. Under SSP585, 75% of the 8,000 simulations display a significantly higher variance (using an F-test) between the 2080-2099 period and the 2001-2020 period; on average, the variance is projected to increase by roughly 300%. This large increase in yield volatility further demonstrates the substantial risk from climate change under a high warming scenario, if growers are not implementing changes in management. We further examine the impact of climate change on almond yield by examining the frequency and severity of yield losses compared with the historical period (2001-2020) under simulations without innovation past 2020 (Fig. 2). By mid-century, 90% of the simulations under either SSP245 or SSP585 project the maximum annual loss to exceed 15 and 20%, respectively, further increasing to at least 25% under SSP245 and 50% under SSP585 by end-of-century. Furthermore, 90% of the simulations show at least 10 and 13 years of loss out of 20 by mid-century under SSP245 and SSP585, respectively, rising to at least 15 years under SSP245 and 19 years under SSP585 over the 2080-2099 period. Meanwhile, by 2100 under

the high warming scenario, 10% of simulations show at least 3 years with 100% loss and all years with a minimum of 35% yield loss. Overall, our analysis shows that climate change will not only impact average yields but will significantly increase the risk of major annual losses and sustained losses over multi-decadal periods. However, Fig. 2 (right panel) shows that California is generally projected to experience frequent but mild yield losses instead of few but severe yield losses.

1.3.3 Spatial distribution of yield change

At the county level (Fig. 4), yield gains from innovation display a large heterogeneity in the northern half of the Central Valley but are more homogeneous in the southern half. Yuba and Sutter counties, two northern counties that are directly adjacent to each other, are modeled to benefit the least from innovation with median yield gains of 0.42 and 0.91 ton/acre by the end of the century with sustained innovation; Colusa county, which is contiguous to Sutter county, however, is projected to experience the largest innovation gains with 2.64 ton/acre. Sustained innovation is projected to result in yield gains between 1 to 2 ton/acre for the other counties by 2100. We note that our approach to represent innovation assumes that historical innovation gains will drive future innovation gains in each county independent of other counties. In reality, there is a state-wide trend in innovation adoption, such as innovations led by Universities or by the California Almond Board, and county-specific adjustments to that trend, based on local farm advisors and historical differences in how large and well-resourced farmers are in one county compared to another. In addition, counties with weak innovation gains coincide with counties that do not have a strong and widespread history of almond production in the last four decades. These counties may have more recent innovation adoption but have not yet fully benefited from these innovations and these adoptions may not have occurred for a sufficient number of years to influence the modeled regressions. For example, given that almond trees begin bearing fruit in the third or fourth year after planting, widespread adoption of new rootstock technology would take at least four years to start to reflect in the data. As a result, innovation trends in this study should not be interpreted as indicating that some counties will continue to fail to adopt innovation just because they have not yet. Instead, they point to heterogeneity in past innovation and the potential for yet-to-be-realized innovation gains in various counties.

The Discrepancy exists in the county-level historical yield changes between LOCA2-Hybrid and

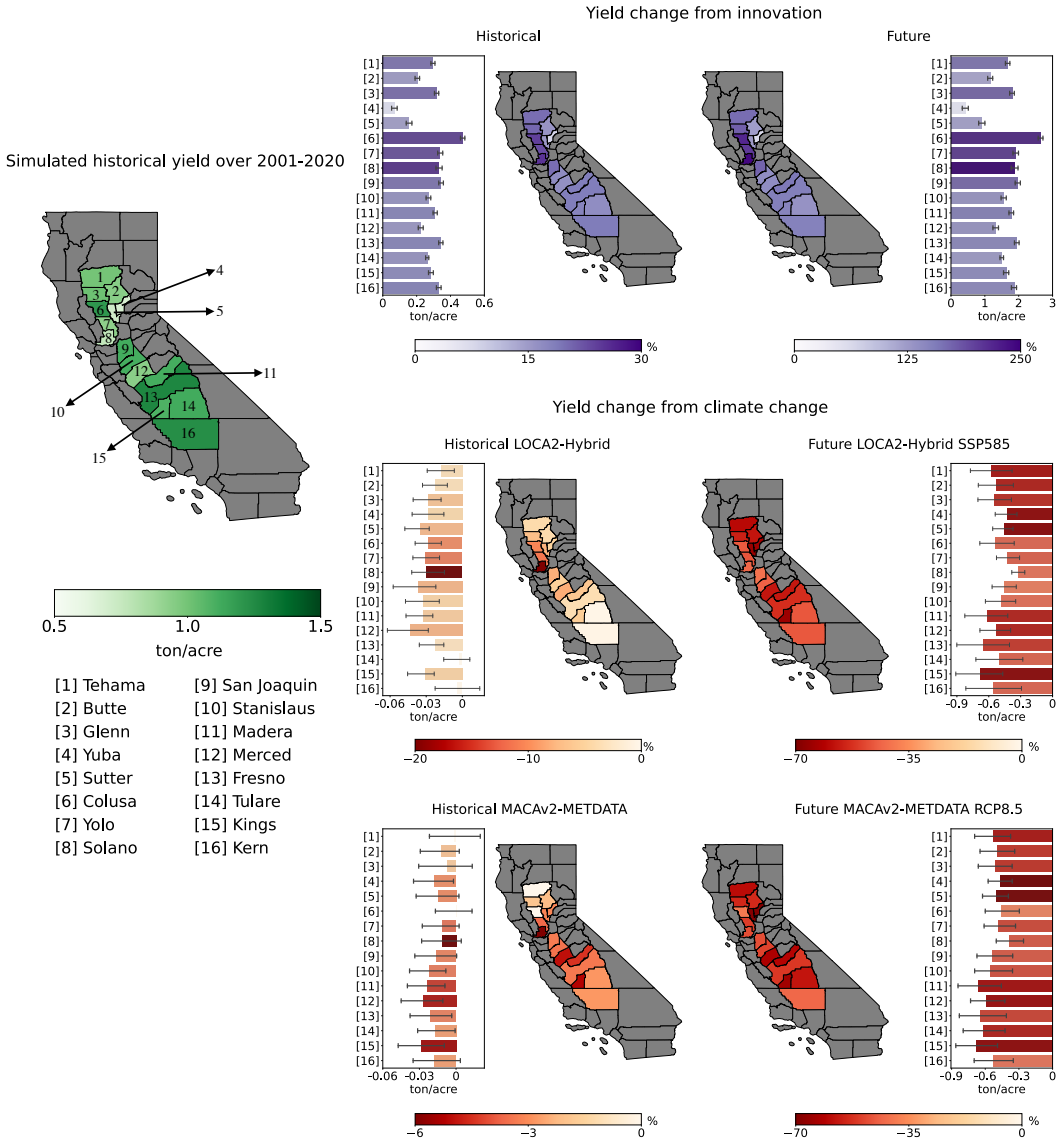

Figure 1.4: California map of county-level historical yields (**green**) and yield changes due to climate change (**red**) and innovation (**purple**). A total of 16 counties in California produce most of the almonds. The county-level historical yields are simulated over the period of 2001-2020. The historical yield changes are calculated by comparing simulated yields over 2001-2020 with those over 1981-2000, and the future yield changes are calculated by comparing simulated yields over 2080-2099 with those over 2001-2020 under high warming scenario (SSP585 for LOCA2-Hybrid and RCP8.5 for MACAv2-METDATA) for those due to climate change, and under sustained innovation scenario for those due to innovation. The **green** and **purple** maps and the color of each bar in bar plots indicate the yield changes in percentage for each county; The bar plots show the yield changes in magnitude with error bars representing ranges between 25 and 75% percentiles for each county.

MACAv2-METDATA ensembles. In the LOCA2-Hybrid ensemble, Tulare and Kern counties, which are located in the south, experience the least impact from climate change with yield losses less than 0.005 ton/acre. Meanwhile, in the MACAv2-METDATA ensemble, the northern counties of Tehama and Colusa show the least climate change impacts. Generally, LOCA2-Hybrid simulations demonstrate a more severe yield loss over the historical period than MACAv2-METDATA simulations. Meanwhile, MACAv2-METDATA simulations show a pattern of climate change impact whereby the southern half of the Central Valley experiences higher climate damages than the northern half over the historical period, a pattern that is less clear in the LOCA2-Hybrid ensemble. However, Solano county experiences the largest yield decline (in percentage) in both ensembles. This shows that both ensembles have local differences in historical trends.

While the county-level historical yield changes show clear differences between the LOCA2-Hybrid and MACAv2-METDATA, the future yield changes between the two ensembles are very similar in terms of magnitudes, percentages, and spatial distributions with median gain losses from climate change ranging between 0.33 and 0.71 ton/acre under SSP585 (LOCA2-Hybrid) and between 0.39 and 0.7 ton/acre under RCP8.5 (MACAv2-METDATA) over the entire Central Valley. Since counties in the northern half of the Central Valley tend to have lower yields over the historical period, this leads to large yield losses associated with climate change in term of percentage for several counties. The county-level analysis also shows large variations in yield changes among counties. For example, future almond yields in Yuba and Sutter are projected to decline by about 65% from climate change under SSP585 by 2100. Meanwhile, Colusa county is projected to be the least impacted by climate change and to gain the most from innovation. By contrast, Merced county is projected to experience moderate innovation gains and suffer larger losses from climate change.

1.3.4 Uncertainty of yield projection

This analysis is subject to several sources of uncertainty, namely internal climate variability, statistical model, climate model, climate scenarios, and innovation scenario uncertainty. By applying the approach from (Hawkins & Sutton, 2009), we partition the sources of uncertainty and present their fractional contributions to the overall uncertainty in yield projections (Fig. 4). As expected, the overall uncertainty keeps increasing with time. Except for the first few years of the future projections, the uncertainty associated with the choice of innovation scenarios made in this study is

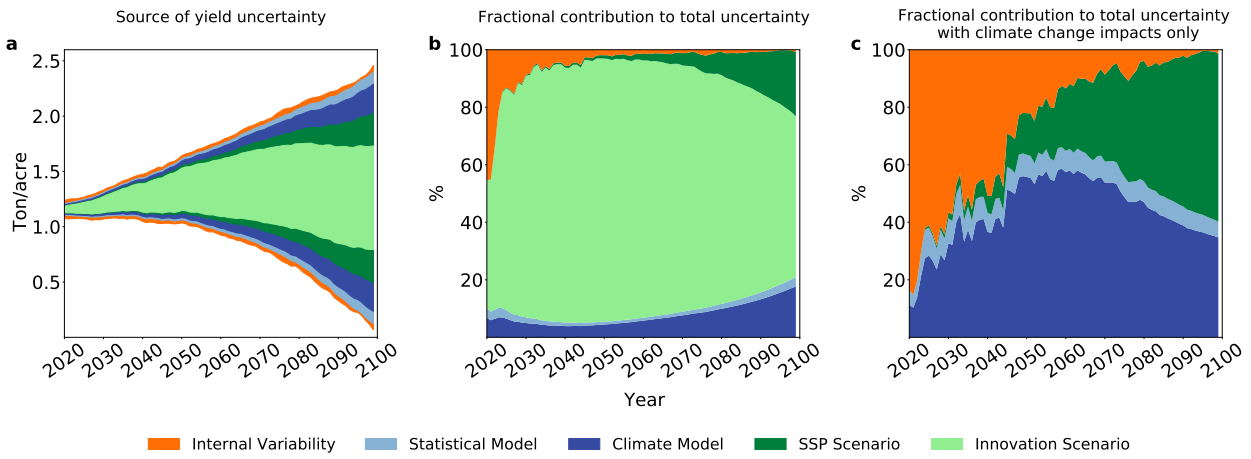

Figure 1.5: Analysis of uncertainty sources within the LOCA2-Hybrid ensemble. The uncertainty of projected yields is partitioned into internal variability, statistical models, climate models, SSP scenarios, and innovation scenarios. **a** shows the time-series of total partitioned uncertainties.**b-c** show the fractional contribution of uncertainties with (**b**) and without innovation scenarios (**c**).

the most important, which reveals that the selection of innovation scenarios largely influences yield simulation results, and remains a major limitation in our ability to predict the fate of the almond industry over the 21st century. The absolute magnitude of the uncertainty due to internal climate variability is nearly constant, and represents the irreducible error in the projections; however, its relative contribution to the total uncertainty decreases continuously over the century. Climate model uncertainty, which represents the differences in yield simulations from using different climate models, is the second dominant contributor to the total uncertainty from 2050 until 2070 when it is overcome by the uncertainty in climate scenarios. Finally, the statistical uncertainty, which represents the differences in yield simulations from using different statistical models, is the smallest source of uncertainty, implying that the 1,000 statistical models trained on different random samplings of the observational record have similar behavior and thus are well-constrained by observations.

The uncertainty analysis of simulations generated from MACAv2-METDATA shows similar partitioning patterns as those from LOCA2-Hybrid (Fig. S2). As 18 GCMs from MACAv2-METDATA were included in the analysis, climate models account for more than half of the uncertainty from the mid-2030s to 2080 without including the innovation scenarios and their maximum fractional contribution reaches nearly 70%. By contrast, 8 GCMs from LOCA2-Hybrid were included in our analysis and their maximum contribution is less than 60%.

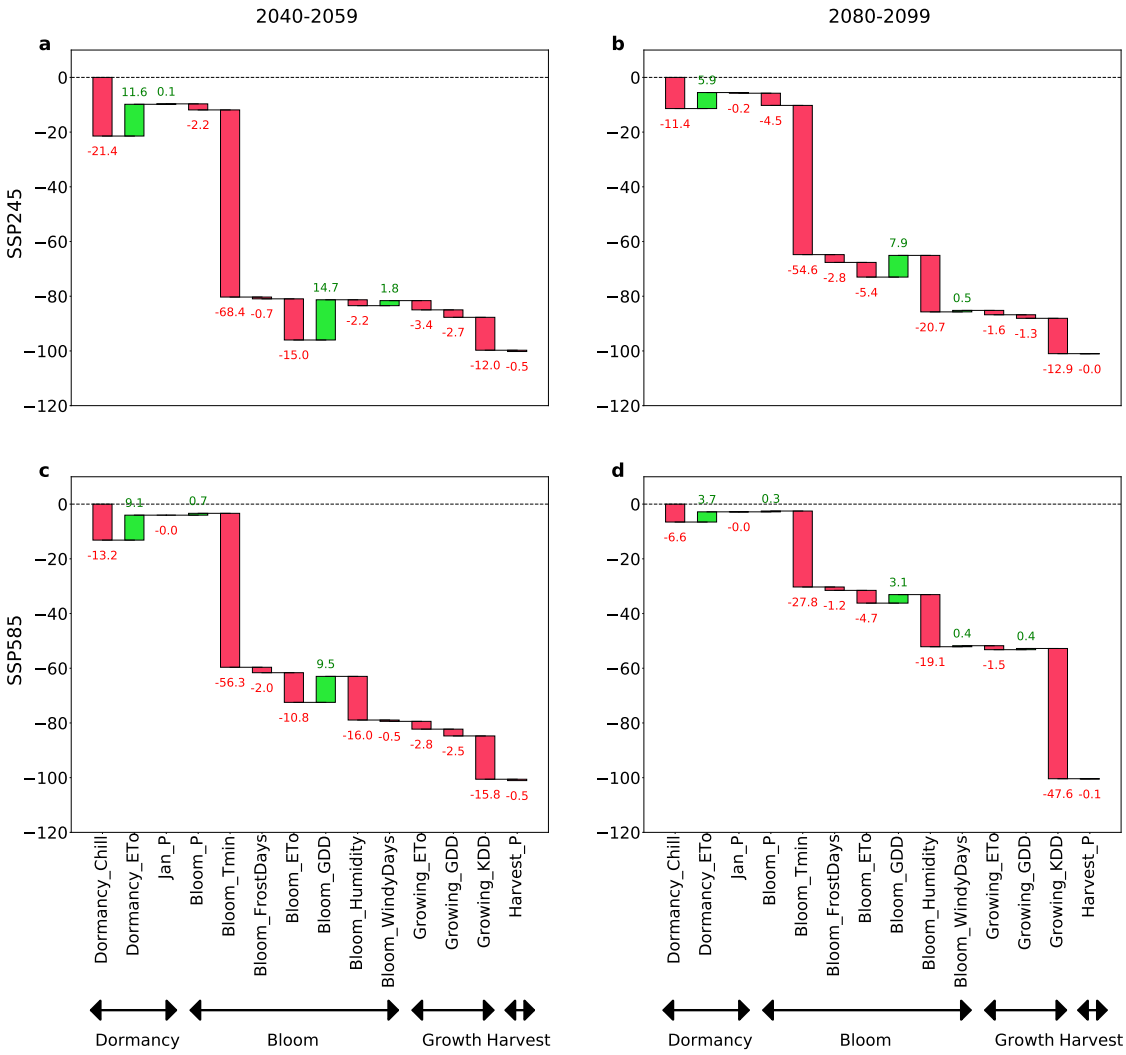

Figure 1.6: Statewide ACI contributions for the LOCA2-Hybrid ensemble. Area-weighted relative contributions (%) of all ACIs are calculated by dividing the magnitude of each ACI's contribution to the magnitude of total yield change caused by climate for 2 climate scenarios (SSP245 and SSP585) and 2 periods (2040-2059 and 2080-2099). Red (green) bars represent negative (positive) contributions.

1.3.5 Relative contributions of ACIs to yield change

The statewide and county-level almond yield projections demonstrate the overall negative effects of future climate change on the California almond industry. Few studies provided actual adaptation strategies to the agricultural industry. Here, we compute the relative contribution of each ACI to the projected almond yield changes to better understand the physiological processes impacted by climate change.

The majority of simulated yield losses under both SSP scenarios is attributed to increases in the daily minimum temperature during the bloom stage (Bloom_Tmin) (Fig. 6). Our results find that changes in Bloom_Tmin account for 68 and 55% of the yield losses at mid-century and

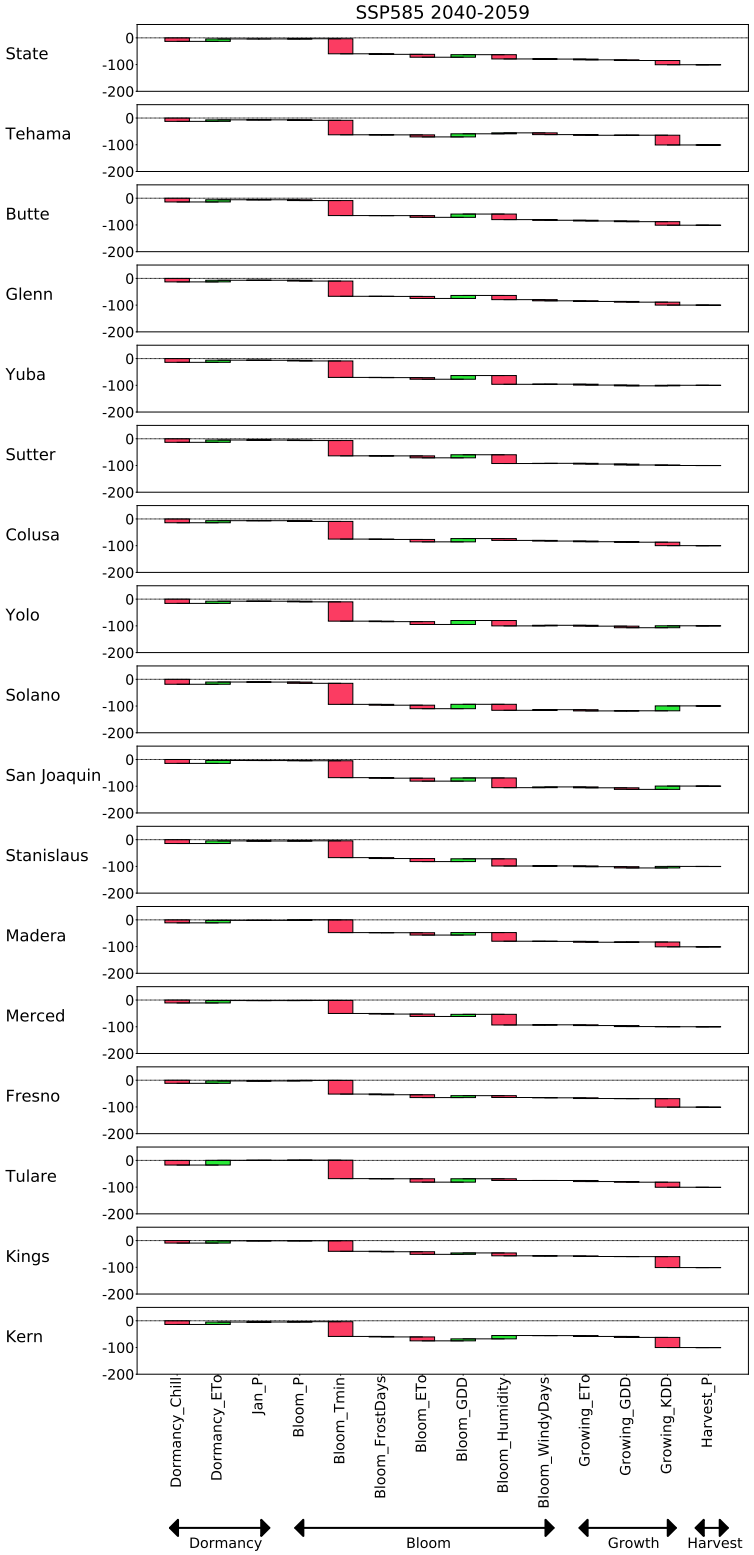

Figure 1.7: County-level ACI contributions for the LOCA2-Hybrid ensemble by 2040-2059 under SSP585. The magnitude of each ACI's contribution for each county is divided by that county's total yield change in magnitude.

end-of-century under SSP245. The relative contributions of Bloom_Tmin decrease slightly under SSP585 but still dominate the yield loss at mid-century being responsible for 56% of yield losses. Increased humidity levels during the bloom stage (Bloom_Humidity) explain 2% (mid-century) and 21% (end-of-century) of yield losses under SSP245, and 16% (mid-century) and 19% (end-of-century) under SSP585. Increases in the exposure to temperatures higher than 35 °C during the growth stage (Growing_KDD) also play a detrimental role by explaining 13 and 48% of yield losses by end-of-century under SSP245 and SSP585 respectively. Meanwhile, other ACIs have noticeable contributions to the simulated yield changes. For example, decreases in accumulated chill hours during the dormant phase (Dormancy_Chill) explain 7 to 12% yield losses. Meanwhile, the increased cumulative exposure to optimal temperatures during the bloom phase (Bloom_GDD) provides small offsetting gains in yield between 3 and 15%. Given that climate change is projected to worsen over the 21st century and between SSP245 and SSP585, the contribution of each ACI to the overall yield change is also evolving with time and warming level. The relative contributions of Bloom_Tmin, Dormancy_Chill, and Bloom_GDD are projected to decrease with time and under increasingly more severe climate change between SSP245 and SSP585. Conversely, Bloom_Humidity and Growing_KDD35 contribute more to the yield losses with more severe climate change, which also occurs in the contribution analysis of MACAv2-METDATA (Fig. S3)

Though the county-level ACI contributions generally follow the same pattern as the state-wide ACI contributions, some variations exist among counties due to differences in local climates (Fig. 7). For example, Bloom_Humidity contributes negatively to yield change in most counties, but Kern and Tehama counties are projected to benefit from changes in Bloom_Humidity during the mid-century period under SSP585. Similarly, Growing_KDD is a damaging ACI for almond yields in all counties except for Solano and San Joaquin counties during the mid-of-century period under SSP585. As a result, this county-level analysis demonstrates that the impacts of climate change on certain phenological development of almonds can vary a lot due to differences in local climates. Similar to the state-wide analysis, the county-level ACI contributions change with different levels of warming (see Fig. S4, S5, and S6). However, they remain consistent and robust between the LOCA2-Hybrid and MACAv2-METDATA ensembles (see Fig. S7, S8, S9, and S10).

1.4 Discussion and conclusion

Perennial crops provide valuable nutrition and economic benefits to society and are a crucial component of the global food supply and agricultural industry. Investigating the climate risks, potential innovation gains, and adaptation opportunities of perennial crops is vital to ensure their resilience. Given the major role of the California almond industry in supporting the state's agriculture and the global supply of almonds as well as its comprehensive data availability, we select California almonds as a case study to evaluate the performance of our modeling approach on exploring the complex interaction between human-driven systems (agriculture and innovation) and natural systems (climate change). In this study, we develop a statistical framework to identify the historical influence of key agro-climate indices (ACI) on specific phenological development stages of almonds, apply it to two large ensembles of future climate projections to assess the future impact of climate change and innovation on almond yields, and assess potential strategies to adapt to climate change. Instead of including seasonally averaged temperature and precipitation (Hong et al., 2020) or only focusing on weather variables in January and February (Lobell et al., 2006), our method provides a direct analysis of the impacts of climate change on key physiological processes driving the phenological development of almonds. Our approach to bring together climate scientists, biometeorology specialists, plant scientists, and agricultural engineers to integrate climate modeling, horticulture and agronomy science, with statistical modeling and machine learning provides important insights into the direct impact of climate change on each physiological growth stage and thus enables the identification of adaptation strategies to minimize future yield losses that current crop modeling studies do not. In comparison, most crop modeling studies provide projections of future crop losses due to climate change but they cannot identify what physiological pathways are responsible for these impacts. We argue that our modeling approach can be applied not only to other almond growing regions, but also to other perennial crops like other nut and fruit trees, and even annual crops, after identifying appropriate ACIs.

We find that climate change could have a substantial impact on future almond yields in California, with projected median yield losses of 17% and 49% by the end of the century for SSP245 and SSP585, respectively, if growers do not change their current management. Future innovations in technology and management practices could potentially offset these adverse effects and even result in

yield gains over the 21st century. We also find that the negative impacts of climate change are more severe in the South of the California Central Valley than in the North, while innovation gains show some strong heterogeneity in northern counties, with overall lower gains compared to the South. The distribution and adoption of agricultural innovations are subject to various factors such as water availability, policy interventions, farm sizes, and crop prices (Caswell & Zilberman, 1985; Feder & Umali, 1993; Pardey et al., 2010). The history of large-scale production of almonds in the South of the Central Valley facilitates the adoption of agricultural innovations in these counties. By contrast, counties that do not have a history of widespread almond production and large-scale farming operations over the last four decades, like Yuba and Sutter counties (USDA National Agricultural Statistics Service, 2021, 2022), have started to adopt innovations in recent years but have yet to benefit from these innovations. However, the prospective introduction and adoption of existing agronomic technology and practices also indicate a large potential of yield increases in these counties, which are not modeled in this study. Large uncertainties exist in our analysis, driven by four major sources: the future innovation scenario (the largest overall source of uncertainty), the internal variability in the climate system (largest in the first decade), the climate model uncertainty (largest during mid-century) and the warming level scenario (largest by end-of-century). The uncertainty partitioning in this study is consistent with analyses of projected temperature and precipitation conducted in previous studies (Hawkins & Sutton, 2009, 2011; Lehner et al., 2020; Monier et al., 2015). The relatively small uncertainty across the 1,000 statistical models demonstrates that the statistical modeling framework developed and used in this study is well-constrained and does not suffer from overfitting, thus providing additional confidence in our approach.

We further identify the specific climatic conditions that are projected to reduce almond yields and the phenological growth stages that are impacted. The current literature modeling the effects of climate change on crop yields often lack details on the impact pathways of climate damages focusing more on estimating yield changes instead. This is true for studies focusing on annual crops (Jägermeyr et al., 2021; Rosenzweig et al., 2014; Wing et al., 2021) and perennial crops (Hong et al., 2020; Lobell & Field, 2011; Medellín-Azuara et al., 2011). By identifying the specific physiological processes and phenological growth stages involved in the yield losses, we aim to provide actionable information for growers, farmers, researchers, funders, and policy-makers to focus their adaptation efforts and identify what adaptation strategies to prioritize. We find that the dominant damaging

ACIs are daily minimum temperature during the bloom stage (Bloom_Tmin), especially under mild warming by mid-century or under the lower warming scenario SSP245, as well as humidity levels during the bloom stage (Bloom_Humidity) and the exposure to temperature higher than 35°C during the growth stage (Growing_KDD), especially under higher warming such as by end-of-century or under the higher warming scenario SSP585. Increases in Bloom_Tmin can lead to earlier and quicker blooming for certain commercial cultivars while the pollenizer cultivars might not bloom yet, which dramatically shortens the overlap of blooming and shorten the cross-pollination period, which would lead to compromised pollination and decreased yields (K. S. Pope et al., 2015). In addition, higher Bloom_Tmin could affect stigma's reception of pollen, pollen tube growth and ovule viability, and thereby damage the overall productivity of almonds (Ortega et al., 2004). The damaging impacts of increasing minimum temperatures during bloom have only been considered in a handful of almond yield modeling studies to date (Lobell & Field, 2011; Lobell et al., 2006). In fact, there has been a large emphasis on examining frost damages during bloom rather than heat damages (Guillamón et al., 2022; Martínez-Gómez et al., 2017; Segura et al., 2017). Higher humidity levels during bloom can also increase the risk of fungal and bacterial diseases, such as blossom and twig blight (Freitas et al., 2023), and can also suppress dehiscence, a key process in pollen release (Gradziel, 2009), and thus result in reduced yields. While the response to extreme heat exposure during the growing season has been extensively studied for annual crops, its underlying mechanisms and impacts are relatively lacking for perennials, especially for crops that are well irrigated and thus can withstand higher temperature levels (Deryng et al., 2014; Lobell et al., 2013; Parker et al., 2020). However, the combined effect of heat stress and water deficit has been shown to affect the photosynthetic capacity, hydraulic system, and carbon efficiency of perennial crops (Carins-Murphy et al., 2023; Haworth et al., 2018). The identification of the specific environmental conditions that are resulting in yield losses under climate damage provide usable information to support future investment in agronomic innovation and to formulate adaptation strategies to ensure limited impacts from climate change on the California almond industry.

Our results suggest there exists a potential synergy between continued innovation yield gains and climate adaptation strategies to counteract the detrimental effect of future climate change. Such strategies can range from the development of resilient almonds through breeding and genetic editing to the improvement of orchard conditions through management practices. However, identifying the

specific damaging climatic conditions as well as the physiological processes and phenological stages affected is crucial to developing almond breeds resilient to climate change through genetic editing (Driedonks et al., 2016; Hedhly, 2011; Kakoulidou et al., 2021). In addition, understanding the timing of climate damages is key to prioritizing adaptation strategies. Since almonds are commonly self-sterile, they require cross-pollination and pollinators like the honeybees, which promotes their genetic variability and facilitates the breeding of new cultivars with favorable traits (Gradziel, 2009), but leaves them vulnerable to warming minimum bloom temperature. Previous almond breeding programs mostly focused on developing and planting late-blooming cultivars to avoid frost damages in Mediterranean regions and pursuing self-fruitfulness to ensure sufficient pollination in California (Gradziel, 2022; Segura et al., 2017). Besides continuing to introduce and invest in early-blooming almond cultivars and self-compatible cultivars with a high level of autogamy, innovation efforts could focus on identifying the genotypes that show resilience to warming during the pollination and breeding new cultivars with such genetic traits to ensure pollination success under climate change (Gradziel et al., 2007; Henselek et al., 2018; Kodad et al., 2009). Orchard managers can try to mitigate the rise in daily minimum temperature in orchards or even shift orchards to cooler areas (Parker & Abatzoglou, 2018). Similarly, potential strategies to counteract the negative impact of higher humidity levels include improved fungal disease management and enhancing almond's genetic resistance to diseases (Adaskaveg et al., 2022; Gradziel & Martínez-Gómez, 2013; Palacio-Bielsa et al., 2017). To account for the intensifying heat stress during the growing stage, optimizing irrigation systems and breeding cultivars with high tolerance to drought and heat are potential solutions (Campalans et al., 2001; Parker et al., 2020). Given the relative impacts of Bloom_Tmin, Bloom_Humidity, and Growing_KDD under different levels of warming, priorities should be given to address pollination resilience in the coming decades, and then focusing on fungal disease resilience and management efforts, as well as heat tolerance in future decades. Adopting adaptation strategies that target specific climate damages, as identified in this study, is more likely to limit harvest failure, stabilize the agricultural market, and ensure the well-being of agricultural worker communities.

While this study provides important insights into the future of almonds in California, we acknowledge several limitations in our modeling framework, which should be addressed in future research. First, the reliance on a statistical modeling framework limits our ability to account for unknown physiological drivers of yield changes. Because almond orchards in California are

well-irrigated, with the reliance on groundwater irrigation during drought periods when surface water availability is scarce, our analysis shows little impacts of water-related agro-climate indices like reference evapotranspiration or precipitation on almond yields. However, with the passage of the Sustainable Groundwater Management Act (SGMA) in 2014, which aims to use and manage groundwater sustainably by 2042, water availability for irrigation may become constrained in the coming decades. As a result, the lack of representation of the impact of future irrigation deficit on almond yield could become a major limitation in our yield projections. This could be addressed in the future by the development of an integrated modeling framework that simulates the impact of climate and socioeconomic changes on water availability for irrigation and its subsequent impacts on almond yields. Such frameworks have been developed for major annual crops, and have shown large yield declines in regions where irrigation is not sustainable (Blanc et al., 2017). It is also important to note that this modeling approach does not account for how extreme conditions leading to crop loss in a particular year might influence yields in subsequent years. This limitation could potentially be addressed by employing a time-dependent model (Z. Zhang et al., 2019) in future efforts. In addition, our statistical modeling framework assumes fixed phenological stages during the study period, thus implying that almonds will not respond to climate change. Instead, a dynamical phenology model could be developed in the future to improve the accuracy of calculated ACIs (K. Pope et al., 2014). Finally, we acknowledge that differences in downscaling and bio-correction can introduce large uncertainty to projections (Lafferty & Sriver, 2023), which is reflected in the county-level historical yield change analysis (Fig. 4). However, our projections of almond yields agree well between LOCA2-Hybrid and MACAv2-METDATA at the state level and county level by mid-century and end-of-century, further validating the statistical modeling framework developed and used in this study and demonstrating the robustness in our analysis.

Since California is the largest almond producer in the world, the future changes in average yields and year-to-year yield fluctuations projected in this study have the potential to affect the global almond market, and thus have large socioeconomic implications. Integrating our results into economic modeling frameworks that can estimate price and economic welfare impacts and supply and demand dynamics would provide additional insights into the fate of California almonds over the 21st century (Medellín-Azuara et al., 2011; Moore et al., 2017; Wing et al., 2015). For example, Medellín-Azuara et al. (2011) uses the Statewide Agricultural Production Model (SWAP) to

examine the changes in price, land use, water supply, and revenue of major agricultural commodities aggregated by types in California in the context of a growing demand and estimates of innovation gains and climate damages from the literature. Our analysis also shows that innovation gains have the potential to shape the future of almonds in California over the next several decades. For this reason, developing a greater understanding of drivers of innovation and their potential impacts on future yields (Chemeris et al., 2022) is key to developing science-based future innovation scenarios, which are a major gap in long-term agricultural climate change impact studies. For example, the production risk and higher crops prices caused by climate change could drive the advancement of agricultural technology and innovation, and facilitate their adoptions by growers (Pardey et al., 2010; Sassenrath et al., 2008; Sunding & Zilberman, 2001). It is also important to better understand the potential implications of future innovation on farming communities; for example, automation in farming can have complex political, social, cultural, and security implications that need to be better understood and examined (Sparrow & Howard, 2021). Furthermore, since our study identifies increases in daily minimum temperatures in February as the leading cause of almond yield declines in the coming decades, we argue that a better understanding of the underlying physiological processes driving the associated yield declines would benefit the almond industry. Similarly, since the relative impact of ACIs on future almond yield losses can change by county, we recommend that additional studies be conducted using more detailed yield information beyond county-level data, such as field studies, to further refine our understanding of the impact of climate change on the phenology of almonds. Finally, we recommend that our approach to bring together climate scientists, biometeorology specialists, plant scientists, and agricultural engineers be applied and evaluated for other almond growing regions, other perennial crops like other nut and fruit trees, and possibly annual crops, after careful identification of appropriate ACIs by experts.

Supplement

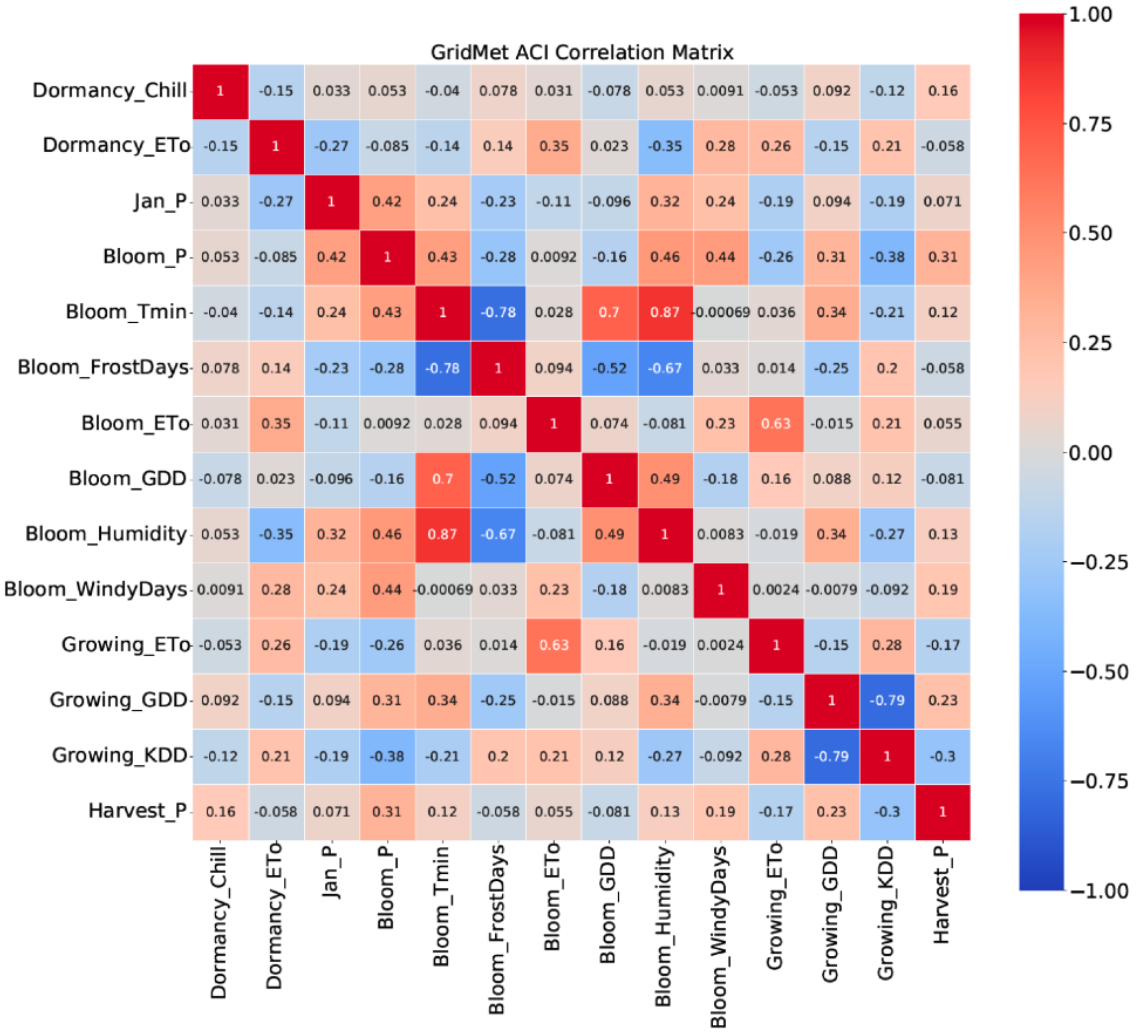

Figure S1: Correlation matrix of ACIs. The value within each cell represents the correlation coefficient between the two corresponding ACIs that are calculated using the GridMet historical meteorological dataset for each county from 1980 to 2020.

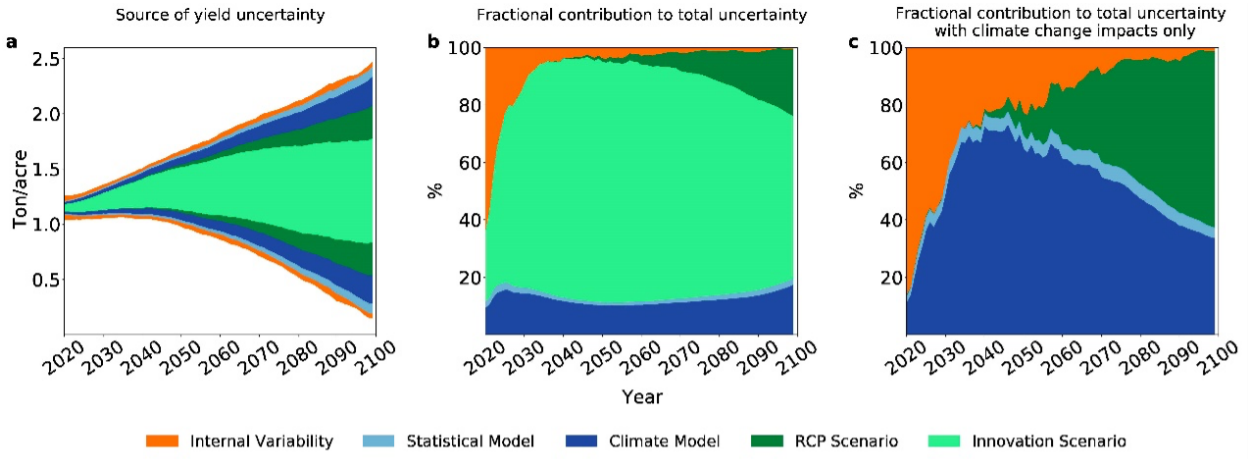

Figure S2: Analysis of uncertainty sources within the MACAv2-METDATA ensemble. See caption of Figure 1.5 for more details.

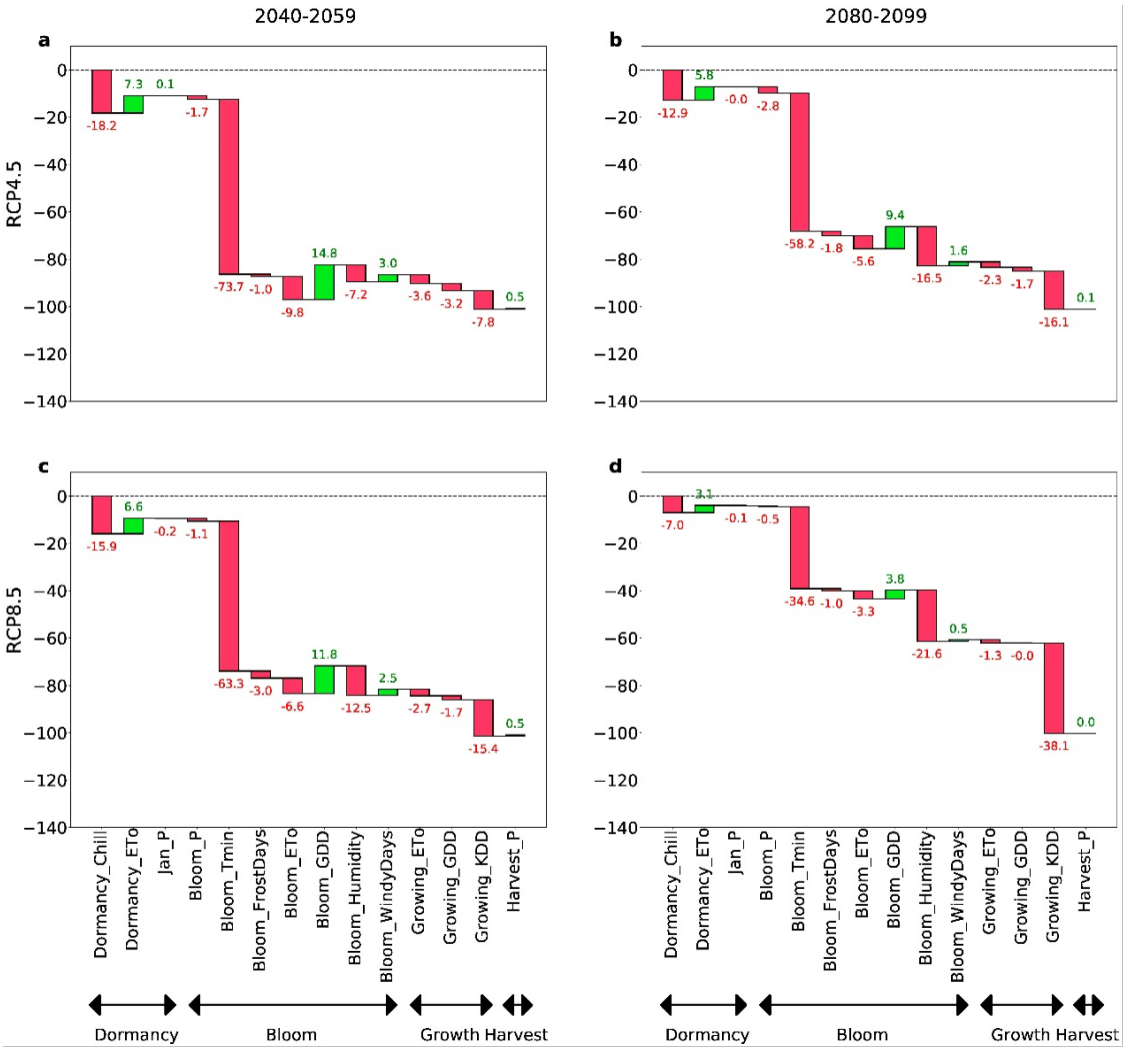

Figure S3: Statewide ACI contributions for the MACAv2-METDATA ensemble. See caption of Figure 1.6 for more details.

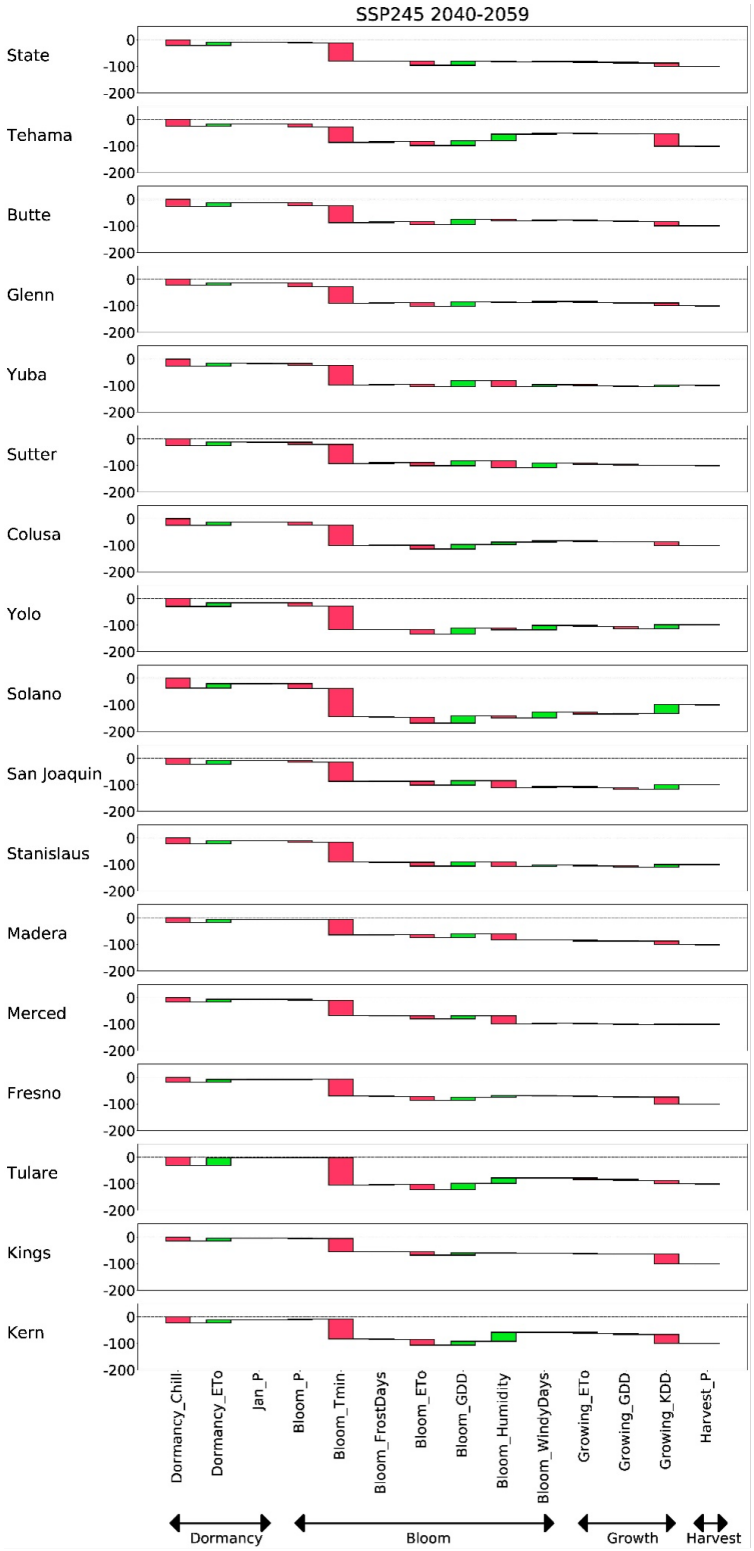

Figure S4: County-level ACI contributions for the LOCA2-Hybrid ensemble by 2040 2059 under SSP245. See caption of Figure 1.7 for more details.

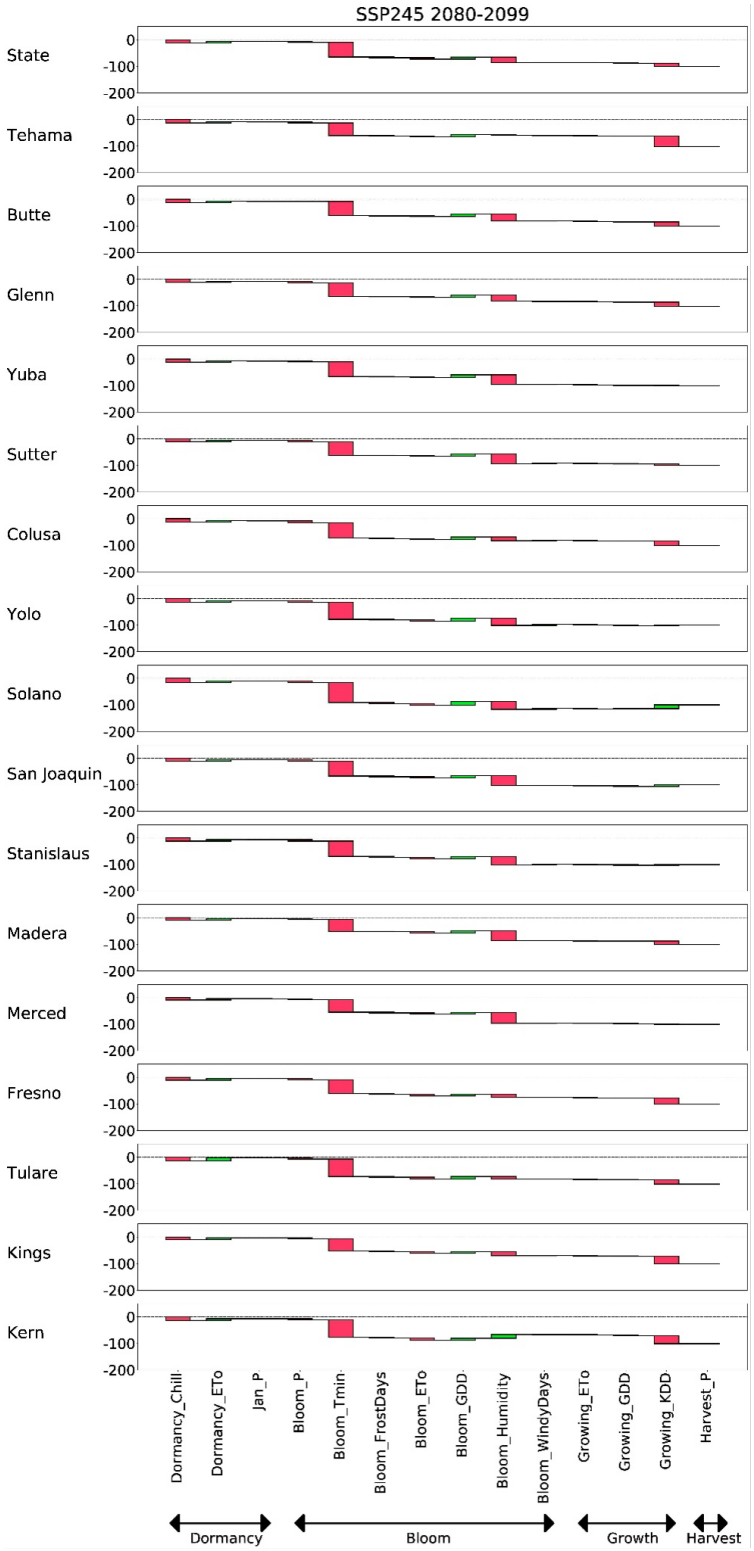

Figure S5: County-level ACI contributions for the LOCA2-Hybrid ensemble by 2080-2099 under SSP245. See caption of Figure 1.7 for more details.

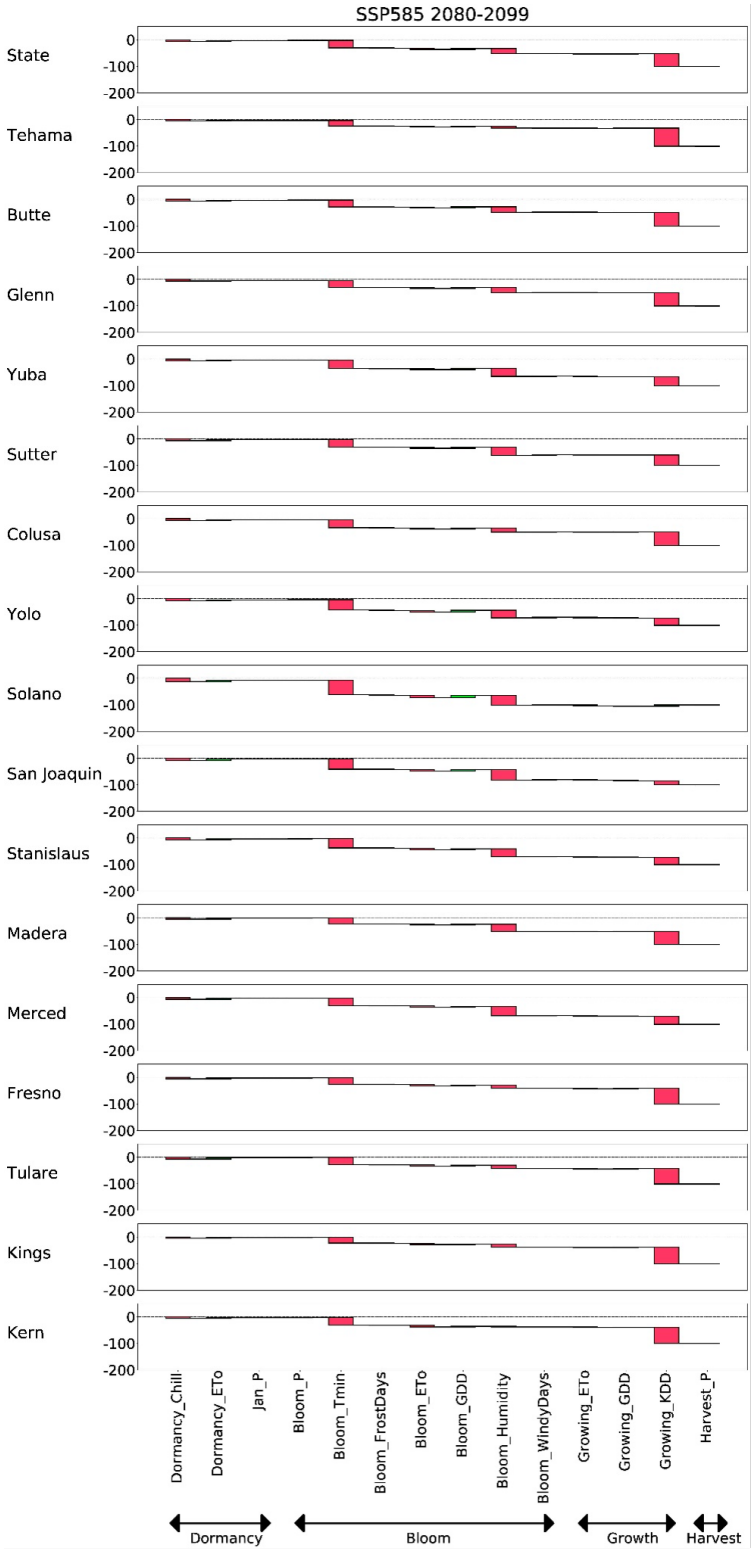

Figure S6: County-level ACI contributions for the LOCA2-Hybrid ensemble by 2080-2099 under SSP585. See caption of Figure 1.7 for more details.

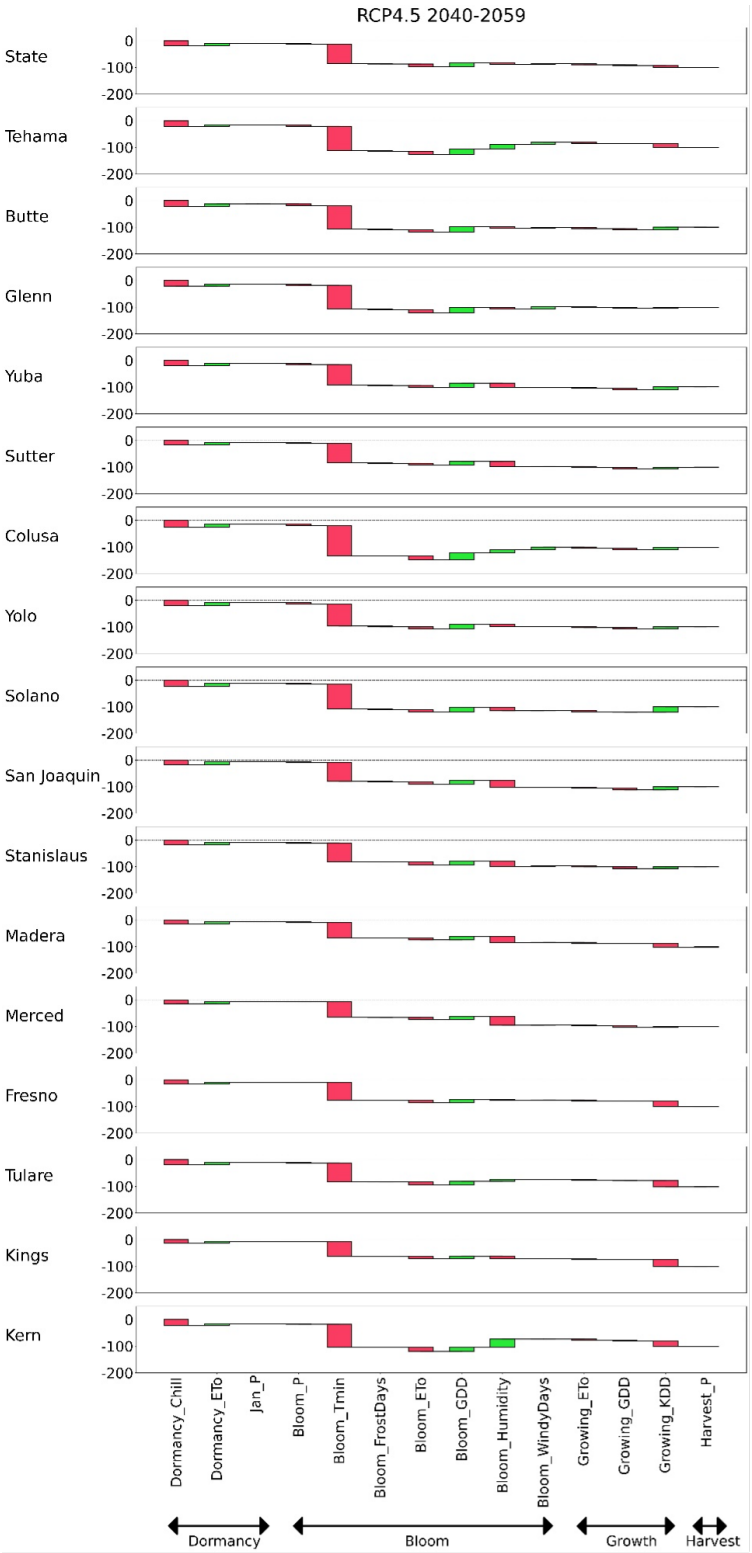

Figure S7: County-level ACI contributions for the MACAv2-METDATA ensemble by 2040-2059 under RCP4.5. See caption of Figure 1.7 for more details.

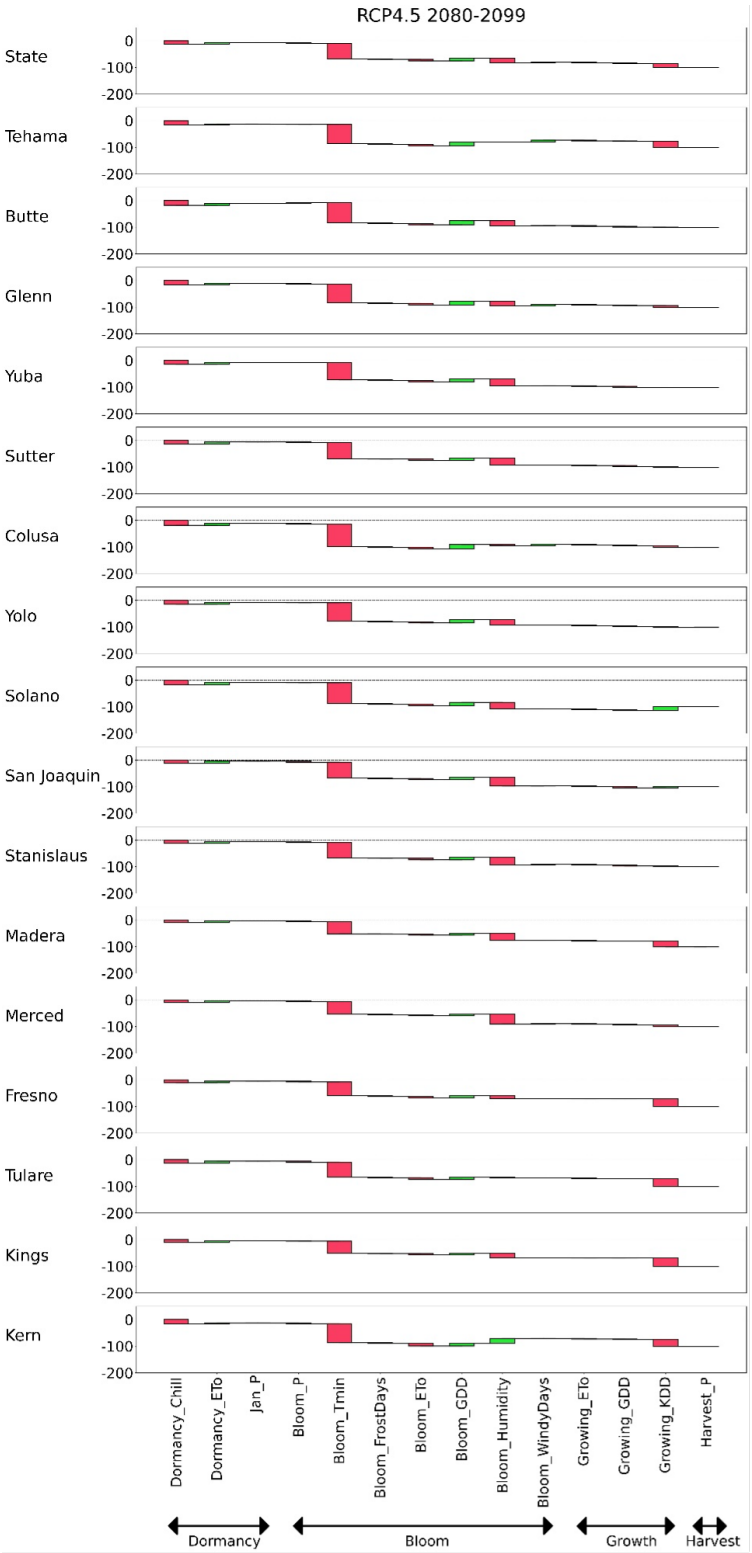

Figure S8: County-level ACI contributions for the MACAv2-METDATA ensemble by 2080-2099 under RCP4.5. See caption of Figure 1.7 for more details.

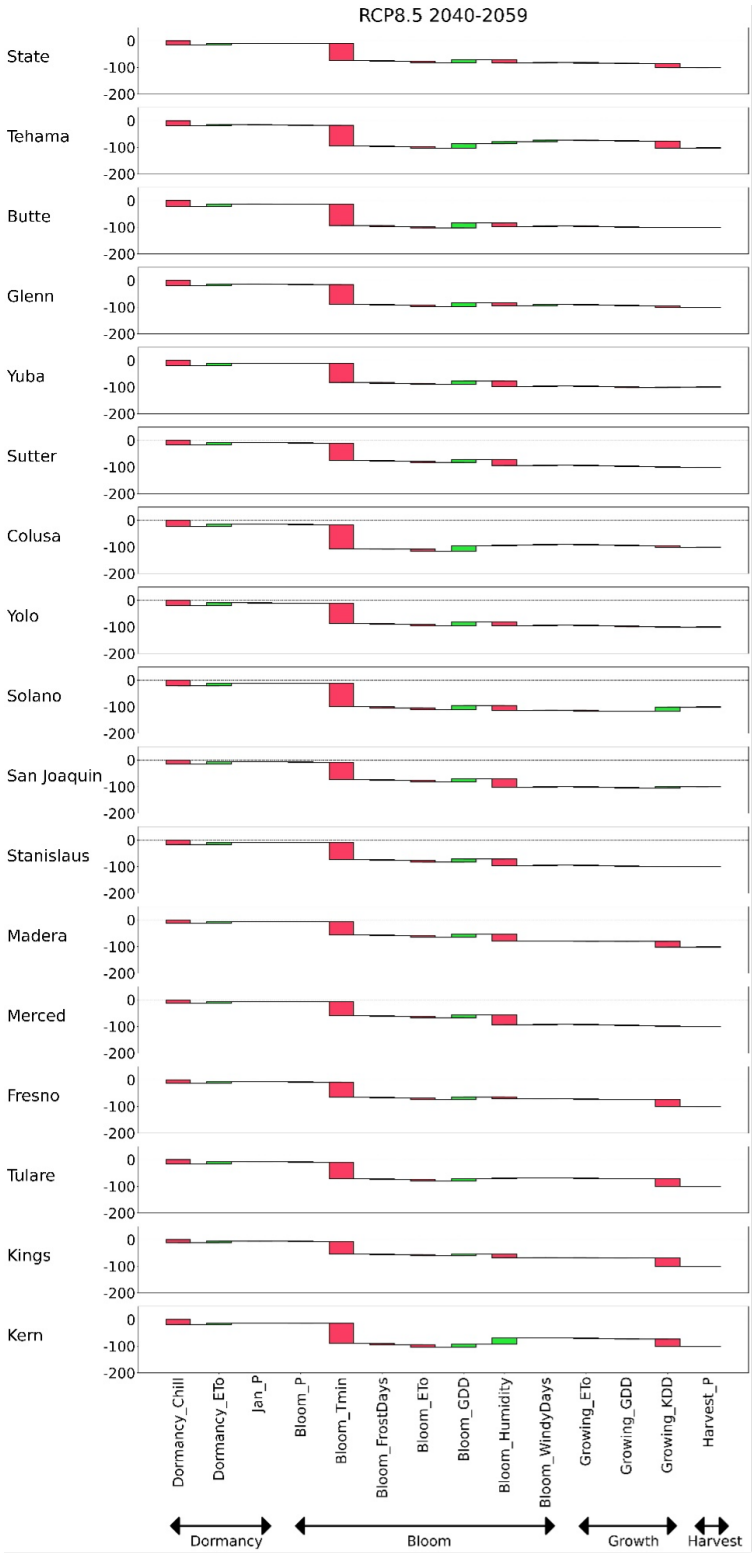

Figure S9: County-level ACI contributions for the MACAv2-METDATA ensemble by 2040-2059 under RCP8.5. See caption of Figure 1.7 for more details.

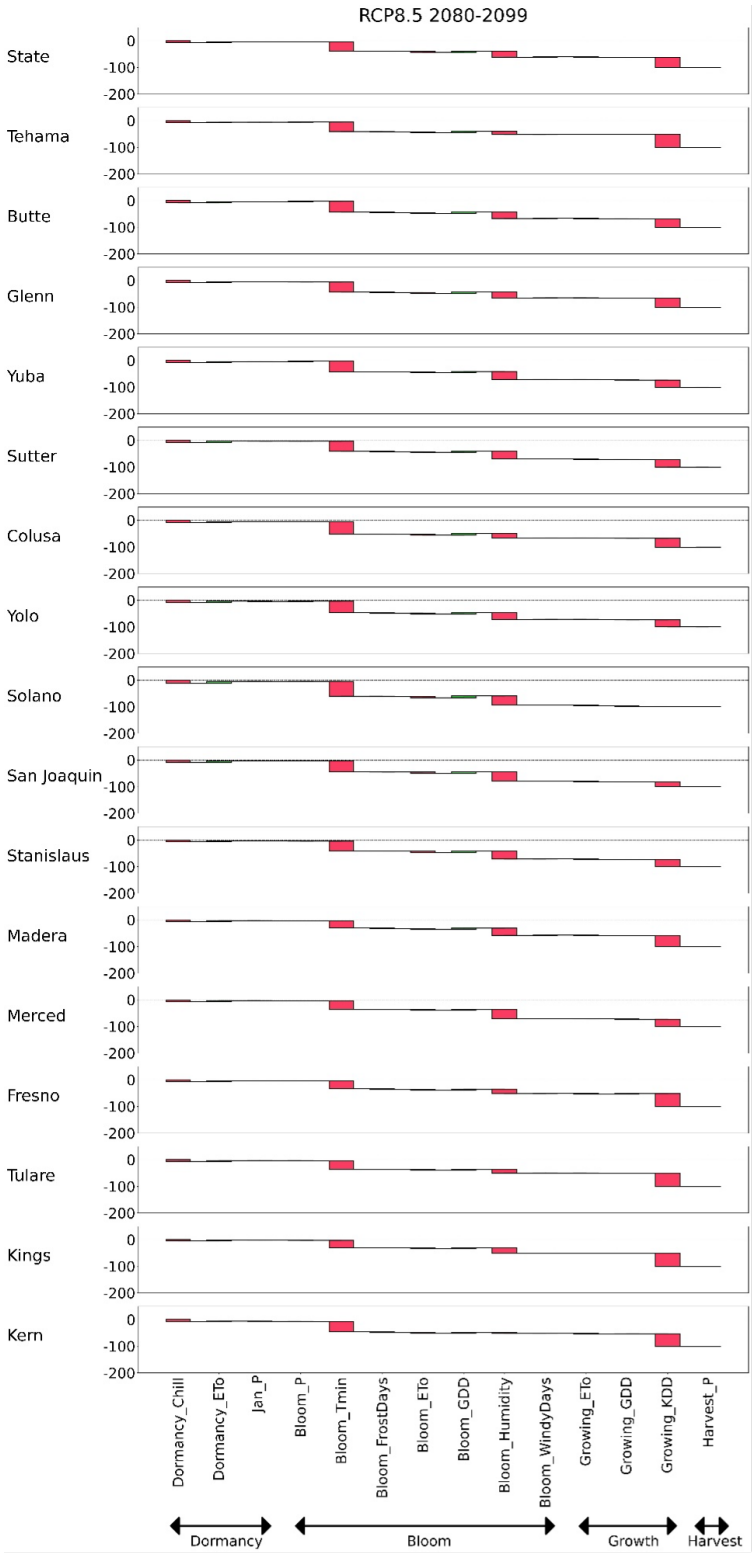

Figure S10: County-level ACI contributions for the MACAv2-METDATA ensemble by 2080-2099 under RCP8.5. See caption of Figure 1.7 for more details.

Chapter 2

Scaling AquaCrop for Continental Applications: A Python-Based, County-Calibrated Framework with a U.S. Maize Case Study

Wu, S., Monier, E. (2025). Scaling AquaCrop for continental applications: A Python-based, county-calibrated framework with a U.S. maize case study. In preparation for the Journal of Advances in Modeling Earth Systems.

Abstract

Process-based crop models are widely applied to investigate environmental impacts on agricultural productivity, optimize agricultural management practices, and support adaptation strategies. AquaCrop, developed by the Food and Agriculture Organization, simulates crop growth and response to water stress and has been applied extensively at the field scale and more sparingly at regional scales. However, most existing parameterizations are based on limited site-specific experiments, raising concerns about their ability to capture spatial and temporal dynamics of maize productivity across large regions. In this study, we present a flexible Python-based implementation of AquaCrop calibrated with two decades of county-level maize yield data of multiple strategically selected counties over the contiguous U.S. Using a global sensitivity analysis, we identify the most influential parameters and calibrated them with yield records from 17 counties representing diverse climate conditions and distinct irrigation managements. This targeted county-level calibration approach substantially improves model performance compared to the default field-scale calibration, enhancing the model's ability to capture both spatial heterogeneity and interannual yield variability. We further compare gridded simulations between county-level calibrated AquaCrop and 14 global crop models participating in GGCMI Phase 1 and show that AquaCrop performs competitively in reproducing yield levels, temporal dynamics, and spatial patterns at state and national scales, with consistent performance across harmonization scenarios. This study demonstrates that regional calibration using county-level yield data can enhance AquaCrop's accuracy, spatial representativeness, and robustness for large-scale applications. More broadly, it highlights AquaCrop's ability to operate at continental scales and establishes a framework that can be extended to other crops and regions.

2.1 Introduction

Food and nutrient insecurity is one of the greatest challenges in the world, with more than 2 billion people experiencing moderate to severe food insecurity in 2023 (FAO, 2024). This crisis is further exacerbated by pressures from overpopulation, climate change, and dwindling water resources (Molotoks et al., 2021; Singh et al., 2017; Wheeler & Von Braun, 2013; Zhao et al., 2017). A growing body of evidence indicates that rising temperatures, shifting precipitation, and more frequent extremes are already depressing yields and increasing volatility across major crops,

including wheat, rice, and maize, with elevated risks of simultaneous production shocks in multiple breadbaskets (Iizumi et al., 2018; Lesk et al., 2016; Lobell et al., 2011; Ray et al., 2015; Tigchelaar et al., 2018). As pressures mount, there is a critical need for tools that translate climate risks into actionable information at the scales where decisions are made.

In this context, crop models serve as pivotal tools to support research, farm-level decisions, and policymaking. Empirical crop models, which develop statistical relationships between observed yield and meteorological data, play an important role in examining the impacts of climate change (Hultgren et al., 2025; Schlenker & Roberts, 2009; Tigchelaar et al., 2018; Wing et al., 2015); however, because they are generally limited to data-rich crops and regions and do not explicitly represent physiological processes or crop management (i.e., soil management, planting dates, irrigation management), they are limited in scope. Meanwhile, process-based models, which explicitly represent the physiological processes of growth as impacted by environmental conditions, are employed to estimate yield gaps, optimize agricultural management practices, evaluate climate change impacts, support crop breeding programs, and develop adaptation strategies (Bo et al., 2024; Chapman, 2008; Chun et al., 2016; Dlamini et al., 2023; Hammer et al., 2006; Kang et al., 2009; M. Kumar, 2016; Reynolds et al., 2018; Tita et al., 2025; Weih et al., 2022). Large coordinated efforts such as the Agricultural Model Intercomparison and Improvement Project (AgMIP, Rosenzweig et al., 2013), a major international collaborative effort to assess and improve global agricultural modeling, and its Global Gridded Crop Model Intercomparison initiative (GGCMI, Elliott et al., 2015; Müller et al., 2017, 2019), which brings together process-based crop modeling teams to conduct coordinated modeling exercises on climate impacts, advance understanding and benchmarking of climate impacts. While these efforts lead to several seminal studies (Franke et al., 2020; Jägermeyr et al., 2021; Rosenzweig et al., 2014), the underlying models are often challenging for stakeholders to use and run at scales that may not always provide decision-relevant information.

One widely used process-based crop model is AquaCrop, which is developed by the United Nations Food and Agriculture Organization and focuses on simulating crop growth and response to water stress (Steduto et al., 2009). At the field scale, AquaCrop is extensively applied for various crops such as maize, rice, wheat, and cotton, to estimate crop productivity and optimize irrigation amount and scheduling across the world (García-Vila & Fereres, 2012; Khasraei & Eslamian, 2024; F. Li et al., 2019; Sharafkhane et al., 2024; Tang et al., 2023). The application of AquaCrop is also

upscaled to the regional or national level in some regions. For example, Lan et al. (2021) employ AquaCrop to estimate the exploitable yield and harvest area gaps for rice in China. Busschaert et al. (2022) apply it to project the net irrigation requirement of C3 field crops for Europe under different climate scenarios. However, AquaCrop is rarely involved in large-scale studies in the U.S., and its national-level application is even more limited. Previous applications of AquaCrop in the U.S. mostly focus on particular locations or small areas (Fazel et al., 2023; D. Kim & Kaluarachchi, 2015; Sandhu & Irmak, 2019). Furthermore, the few U.S. regional-scale AquaCrop studies to date rely solely on parameters from the default field-scale calibration or on site-specific calibration that may not represent the most accurate model at regional scales (Fazel et al., 2023; Lamsal & Marston, 2025). Existing regional AquaCrop implementations extend the official AquaCrop release by automating input-file generation and streamlining large-scale simulation workflows (de Roos et al., 2021; Lorite et al., 2013). However, they usually lack flexibility in model modification, efficient calibration, data assimilation, and integration with machine learning or artificial intelligence (Chen et al., 2023; Williamson et al., 2023).

In this study, we develop a gridded and Python-based version of AquaCrop and apply an advanced calibration approach using county-level yield data to develop a version of AquaCrop that can run at continental scale. We select maize in the U.S. as a case study because maize is the most extensively cultivated and consumed cereal globally, with roughly 200 million hectares across tropical, subtropical, and temperate zones (Erenstein et al., 2022; FAO, 2024; García-Lara & Serna-Saldivar, 2019), and because the U.S. leads maize production by volume and economic value with almost 30% of the global maize production and over 70 billion dollars in economic value in 2023 (FAO, 2024; USDA NASS, 2025). We implement a version of AquaCrop in Python to enhance flexibility and scalability of the model, and to enable seamless integration with various datasets, visualization tools, and other Python-based models (Chen et al., 2023; Kraft et al., 2010). We use SpotPy (Houska et al., 2015), a Python package offering an extensive set of calibration algorithms, to calibrate AquaCrop in parallel, which dramatically shortens computational time and is readily adopted by other practitioners. The first published Python-based AquaCrop, AquaCrop-OSPy (Kelly & Foster, 2021), represents a significant advance in Python implementations of process-based crop models. However, our analysis reveals that it still exhibits minor discrepancies compared to the official FAO release, which can compound over multi-year simulations. In contrast, our gridded

Python-based implementation produces outputs virtually identical to the official AquaCrop version, thereby improving accuracy in long-term simulations.

By applying advanced calibration techniques on county-level yield data from USDA (USDA NASS, 2023), our approach significantly improves the accuracy and applicability of AquaCrop for large-scale crop simulations. We leverage high-resolution spatial data on climate, soil, and management practices to simulate both county-level and gridded maize yields for the contiguous United States. We evaluate model performance against county-level yield and demonstrate that our calibration approach substantially enhances the ability to capture both spatial and temporal dynamics of observed maize yields under both rainfed and irrigated conditions, compared to the default field-scale calibration. We further run gridded simulations with our county-level calibrated model following the GGCMI Phase 1 protocol and demonstrate that our version of AquaCrop performs well compared to the GGCMI process-based crop models. Overall, this work provides a transparent and reproducible pathway to develop a version of AquaCrop that can be used over the entire U.S. and thus provide decision-relevant county-scale simulation of crop yield, supporting efforts to enhance agricultural resilience and inform decision-making under changing environmental conditions.

2.2 Materials and Methods

2.2.1 Model Description

The Food and Agriculture Organization (FAO) develops AquaCrop to simulate crop growth and responses to water stress with relatively few data and parameter requirements to balance the simplicity, robustness, and accuracy of the model. AquaCrop provides a straightforward and simplified running procedure, allowing governmental agencies, consulting engineers, and other agricultural stakeholders to model crop growth conveniently (Steduto et al., 2009). The crops simulated by AquaCrop include maize, potato, quinoa, rice, soybean, sugar beet, sunflower, wheat, barley, sugar cane, sorghum, and tef. Various studies calibrate AquaCrop for additional crops, including lettuce and table grape (Amirouche et al., 2021; Er-Raki et al., 2021), demonstrating the applicability of the model to a wide range of crops including cereals, vegetables, fruits, legumes, oil crops, and sugar crops. AquaCrop simulates crop growth (i.e., canopy cover, CC), transpiration

(Tr), aboveground biomass production (B), and yield (Y) at a daily step (Steduto et al., 2009). Several studies show that AquaCrop performs comparably to, and sometimes better than, other widely used process-based models, highlighting its potential for large-scale application (Abedinpour, 2021; Battisti et al., 2017; Feleke et al., 2021).

Soil Water Balance

AquaCrop divides the soil profile into 12 compartments with a depth of 0.1 meters by default. If the crop's effective rooting depth exceeds 1.2 meters, the depth of each compartment is adjusted to cover the root zone. Each soil compartment has its own hydrological characteristics based on the soil profile inputs. At each time step (day), AquaCrop estimates the soil water balance by simulating redistribution between soil compartments, deep percolation, surface runoff, capillary rise, soil evaporation, and crop transpiration. The redistribution of water within the soil is described by the drainage function:

$$\frac{\Delta\theta_i}{\Delta t} = 0.0866\ (K_{SAT,i})^{0.35}\ (\theta_{SAT} - \theta_{FC})\frac{e^{\theta_i - \theta_{FC}} - 1}{e^{\theta_{SAT} - \theta_{FC}} - 1} \tag{2.1}$$

where $\frac{\Delta\theta_i}{\Delta t}$ represents the daily change in water content at depth i; K_{SAT} is the hydraulic conductivity at saturation; θ_i, θ_{SAT}, and θ_{FC} are the actual soil water content at depth i, soil water content at saturation, and soil water content at field capacity, respectively. The drainage ability (i.e., $\frac{\Delta\theta}{\Delta t}$) of each compartment is compared with its overlying one to decide if the compartment keeps the drained water or transfers the water to the next compartment. At the bottom of the soil profile, the net cumulative drainage becomes deep percolation loss. The surface runoff is given by:

$$RO = \frac{(P - I_a)^2}{P + S - I_a} \tag{2.2a}$$

$$S = 254\ (\frac{100}{CN} - 1) \tag{2.2b}$$

where RO is the amount of surface runoff loss; P is the precipitation amount; S refers to the potential maximum soil water retention, which is a function of soil curve number (CN); I_a represents the amount of water that can infiltrate before runoff occurs. With a shallow groundwater table, the

capillary rise is estimated by the following equation:

$$CR = e^{\frac{len(z)-b}{a}} \tag{2.3}$$

where CR is the amount of water from the shallow groundwater; z is the depth of the water table; a and b refer to soil-specific parameters used to describe the capillary rise. The evaporation from soil is calculated in two stages: the energy-limiting stage and the falling-rate stage. The former occurs when precipitation or irrigation wets the soil surface layer and evaporation is determined by the energy available for soil evaporation. Once water is fully evaporated from the surface layer, evaporation enters the falling-rate stage, where water flows upward from the subsurface layer. In this stage, evaporation depends on both available energy and soil hydraulic properties. As soil water content decreases, the ability to transfer water to the surface diminishes, causing evaporation rates to decline over time.

$$E_{Energy\ Limiting} = (1 - (1.72\ CC - CC^2 + 0.3\ CC^3))\ Ke_x\ ET_o \tag{2.4a}$$

$$E_{Falling\ Rate} = K_r(1 - (1.72\ CC - CC^2 + 0.3\ CC^3))\ Ke_x\ ET_o \tag{2.4b}$$

where E is the amount of evaporated water; CC represents the canopy cover; Ke_x refers to the maximum soil evaporation coefficient for a fully wet and unshaded surface; K_r is the evaporation reduction coefficient, which decreases with declining water content in the subsurface layer; ET_o is the reference evapotranspiration, calculated through the FAO Penman-Monteith equation (R. G. Allen et al., 1998):

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma\frac{900}{T_{mean}+273}u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \tag{2.5a}$$

$$e_a = 0.6108\ e^{\frac{17.27T_{min}}{T_{min}+237.3}} \tag{2.5b}$$

where Δ refers to the slope of the vapor pressure curve and γ refers to the psychrometric constant; R_n represents the net radiation at the crop surface and G is the soil heat flux density;

T_{mean} and T_{min} are the daily mean and minimum temperatures, respectively; u_2 refers to the wind speed at 2 meters height; e_s is the saturation vapor pressure. Since the AgMERRA climate dataset (Ruane et al., 2015) only provides relative humidity at the maximum temperature, we consider the vapor pressure at the minimum temperature as the actual vapor pressure (e_a) by assuming the dewpoint temperature is close to the minimum temperature (R. G. Allen et al., 1998).

Growing Degree Days

The fraction of soil surface covered by the canopy (CC) is calculated as a function of growing degree days (GDD) to indicate the growth of the crop, instead of relying on the leaf area index (LAI). The AquaCrop reference manual (Raes et al., 2023b) provides three methods to calculate GDD, and we use the default method for maize (method 2):

$$GDD = \sum T_{avg} - T_{base} \tag{2.6a}$$

$$T_{avg} = \frac{T_{max} + T_{min}}{2} \tag{2.6b}$$

where daily maximum temperature ($T_{\max}$) and minimum temperature ($T_{\min}$) are constrained to an upper threshold and the base temperature.

Canopy Cover

In a stress-free growing season, CC typically experiences three stages: growth, plateau, and decay. The growth of CC is simulated by two equations:

$$CC = CC_o \; e^{t \; CGC} \quad \text{if} \quad CC \leq CC_x/2 \tag{2.7a}$$

$$CC = CC_x \; - \; 0.25 \, \frac{(CC_x)^2}{CC_o} \, e^{-\,t\,CGC} \tag{2.7b}$$

where CC_o and CC_x represent, respectively, the initial canopy cover and maximum canopy cover; t is the accumulated GDD since the planting date and CGC refers to the canopy growth coefficient. After the growth stage, CC stays at CC_x and begins to decline until senescence. The decline of CC

is described by the equation:

$$CC = CC_x\Big[1 - 0.05\big(e^{\frac{3.33\,CDC}{CC_x+2.29}t} - 1\big)\Big] \tag{2.7c}$$

where t is the accumulated GDD since senescence starts and CDC refers to the canopy decline coefficient.

Transpiration

The crop transpiration (Tr) is calculated by multiplying the reference evapotranspiration (ETo) by the crop transpiration coefficient (Kc_{Tr}):

$$Tr = Kc_{Tr}\ ETo \tag{2.8a}$$

The crop transpiration coefficient (Kc_{Tr}) is calculated as follows:

$$Kc_{Tr} = (1.72\ CC\ - CC^2\ + 0.3\ CC^3)\ Kc_{Tr,x} \tag{2.8b}$$

where $Kc_{Tr,x}$ is the crop coefficient for maximum crop transpiration. $Kc_{Tr,x}$ also responds to canopy aging and CO_2 elevation, and starts to reduce when senescence is triggered, as described in the following equations:

$$Kc_{Tr,x,adj} = Kc_{Tr,x}\ - (t-5)\ f_{age}\ CC_x \tag{2.8c}$$

$$Kc_{Tr,x,CO_2,adj} = Kc_{Tr,x,adj}\ \ (1 - 0.05\ \frac{CO_2 - 369.41}{550 - 369.41}) \tag{2.8d}$$

$$Kc_{Tr,x,sen} = Kc_{Tr,x,CO_2,adj}\ (\frac{CC}{CC_x})^a \tag{2.8e}$$

where t is the time in days after CC reaches CC_x; f_{age} is a crop-specific parameter that indicates the reduction of $Kc_{Tr,x}$ in response to canopy aging; a refers to a model parameter that controls the extent of decline in transpiration efficiency.

Harvest Index and Biomass

The harvest index (HI) is the ratio of the dry grain yield to the total above-ground biomass. During the yield formation period, HI increases following a logistic growth pattern.

$$HI_i = \frac{HI_{ini}\ HI_o}{HI_{ini} + (HI_o - HI_{ini})\ e^{-(HIGC)\ t}} \tag{2.9}$$

In this equation, i indicates the day after planting and t indicates the day after flowering; $HIGC$ refers to the growth coefficient of HI. The initial harvest index, HI_i, is set to 0.01 in the model, while the reference harvest index, HI_o, is a crop-specific parameter. To obtain the above-ground biomass (B), water productivity (WP) is normalized for CO_2 concentration and multiplied by the accumulation of Tr/ETo:

$$B = WP^* \quad \sum \frac{Tr}{ETo} \tag{2.10}$$

The normalized water productivity (WP^*) needs to be adjusted if the CO_2 concentration is different from the reference value (369.41 ppm). Finally, the harvest index (HI) is used to partition B into the attainable yield:

$$Yield = HI * B \tag{2.11}$$

Water Stress and Temperature Stress

Water stress is the most important driving force for attainable yield and is deeply involved in the simulation process of AquaCrop. Water stress is determined by comparing the depletion (water shortage versus field capacity) with upper and lower water level thresholds, which are distinct for canopy expansion, pollination, canopy senescence, and stomatal closure. Water stress slows canopy growth, accelerates canopy decline, and can even trigger early senescence if severe enough. Stress-induced stomatal closure decreases transpiration (Tr) and limits crop root zone expansion. Additionally, HI is reduced if water stress affects pollination or stomatal conductance (Steduto et al., 2009). To represent the impacts of water shortage on these crop physiological processes, AquaCrop employs four water stress coefficients: Ks_{exp} for canopy expansion, Ks_{pol} for pollination, Ks_{sen} for canopy senescence, and Ks_{sto} for stomatal closure. These coefficients are calculated

through the following equations:

$$K_s = 1 - \frac{e^{S_{rel} * f_{shape}} - 1}{e^{f_{shape}} - 1} \tag{2.12a}$$

$$K_s = 1 - S_{rel} \tag{2.12b}$$

$$S_{rel} = \frac{\frac{Dr}{TAW} - P_{up}}{P_{low} - P_{up}} \tag{2.12c}$$

where S_{rel} is the relative stress level; f_{shape} is the shape factor that determines the convex response curve. Each coefficient (Ks_{exp}, Ks_{sen}, and Ks_{sto}) has its distinct f_{shape}. The response of Ks_{pol} to S_{rel} is assumed to be a linear relationship (Eqn. 11b). TAW represents the total available water (i.e., the difference between field capacity and permanent wilting point) and Dr is the difference between field capacity and the actual water content in the root zone. P_{up} refers to the fraction of TAW at which water stress starts to affect crop growth. P_{lower} refers to the fraction of TAW when water stress is at its full strength (i.e., $Ks = 0$). When simulating the water stress effects on canopy cover, CGC is adjusted by Ks_{exp} during the growth stage and CDC is adjusted by Ks_{sen} when early senescence is triggered.

$$CGC_{adj} = Ks_{exp}\ CGC \quad \text{when} \quad CC > 1.25\ CC_o \tag{2.13}$$

$$CDC_{adj} = (1 - Ks_{sen}^{8})\ CDC \quad \text{when} \quad CC > CC_o \tag{2.14}$$

AquaCrop disregards the influence of water stress on canopy cover until germination ends (i.e., $CC > 1.25\ CC_o$) due to the assumed protection of seedlings, which is a common crop modeling practice. AquaCrop allows crop recovery if early senescence occurs by assuming the crop enters a dormant period during which CC declines at a slower pace. The CC during the dormant period is simulated by the following equation:

$$CC_i = CC_o + (\frac{1 - \sum ETo_i}{L_{dormant}})\ (CC_{dormant} - CC_o) \tag{2.15}$$

where subscript i indicates the day since the crop enters the dormant period and $L_{dormant}$ is the accumulation of ETo over the entire dormant period. As long as CC does not decline to the initial canopy cover (CC_o), canopy expansion can continue when sufficient water is available. When estimating the response of transpiration to water stress, AquaCrop considers the effects of water deficit (Ks_{sto}), waterlogging (Ks_{aer}), and cold stress ($Ks_{cold,tr}$).

$$Tr_{adj} = Tr \ \min(Ks_{sto}, Ks_{aer}) \ Ks_{cold,tr} \tag{2.16}$$

The expansion of the crop root is also subject to Ks_{sto}, as described in the following equations:

$$dZ_{adj} = Ks_{sto} \ dZ \quad \text{when} \quad f_{shape,sto,root} = 0 \tag{2.17a}$$

$$dZ_{adj} = dZ \ \frac{e^{Ks_{sto} \ f_{shape,sto,root}} - 1}{e^{f_{shape}} - 1} \quad \text{when} \quad -8 \leq f_{shape,sto,root} < 1 \tag{2.17b}$$

where $f_{shape,sto,root}$ is the shape factor that describes how the root expansion rate responds to stomatal closure. HI is also adjusted in response to a series of effects from pollination failure, reduced transpiration, and inhibited vegetative growth. Water stress occurring before the yield formation period can potentially lead to an increase in HI due to the saved energy in vegetative growth. The coefficient to quantify this positive adjustment, f_{ante}, is calculated by the following equations:

$$f_{ante} = 1 + \frac{1 + sin((1.5 - Ratio_{low})\pi)}{2} \ \frac{\Delta HI_{ante}}{100} \quad \text{when } B_{r,low} \leq B_{rel} < B_{r,top} \tag{2.18a}$$

$$0 \leq Ratio_{low} = \frac{B_{rel} - B_{r,low}}{B_{r,top} - B_{r,low}} \leq 1 \tag{2.18b}$$

$$f_{ante} = 1 + \frac{1 + sin((0.5 + Ratio_{up})\pi)}{2} \ \frac{\Delta HI_{ante}}{100} \quad \text{when } B_{r,top} \leq B_{rel} < B_{r,up} \tag{2.19a}$$

$$0 \leq Ratio_{up} = \frac{B_{rel} - B_{r,top}}{B_{r,up} - B_{top}} \leq 1 \tag{2.19b}$$

where ΔHI_{ante} is the maximum allowable increase of HI_o; B_{rel} refers to the ratio between the actual biomass and the potential biomass at the start of flowering; $B_{r,low}$ and $B_{r,up}$ are the lower and upper limits of the B_{rel} range affecting HI_o, respectively; $B_{r,top}$ indicates the value of B_{rel} at which f_{ante} has the strongest effect. During flowering, the pollination process is susceptible to adverse effects from both water stress and temperature extremes (cold and heat stress), which decrease HI, as described in the following equations:

$$HI_{adj} = \sum_{1}^{j} \ (min(Ks_{pol,j}, Ks_{cold,pol,j}, Ks_{heat,pol,j})(1 + \frac{f_{excess}}{100}) \ F_j \ HI_o) \tag{2.20a}$$

$$F_j = \frac{100}{L_{flowering}} \bar{f_j} \tag{2.20b}$$

where j indicates the number of days since the beginning of flowering; f_{excess} is a crop-specific parameter that represents the overproduction of flowers; $L_{flowering}$ refers to the total length of the flowering period and f_j is the average fraction of flowers at day j; $Ks_{cold,pol}$ and $Ks_{heat,pol}$ represent the cold and heat stress coefficients for crop pollination, respectively. Within the build-up period of HI, AquaCrop accounts for both the positive effect from inhibited leaf expansion ($f_{post,positive}$) and the negative effect from reduced transpiration ($f_{post,negative}$) on HI.

$$f_{post,positive} = 1 + \frac{\sum_{i=1}^{n_{exp}} \frac{1 - Ks_{exp,i}}{a}}{n_{exp}} \tag{2.21a}$$

$$f_{post,negative} = \frac{\sum_{i=1}^{n_{yield}} \left(\sqrt[10]{Ks_{sto,i}} \ (1 - \frac{1 - Ks_{sto,i}}{b}) \right)}{n_{yield}} \tag{2.21b}$$

$$f_{post} = (\frac{n_{exp} \ f_{post,positive} + (n_{yield} - n_{exp})}{n_{yield}}) \ f_{post,negative} \tag{2.21c}$$

where n_{exp} and n_{yield} are the total lengths of the period when vegetative growth is still possible after the start of flowering and the period for building up HI; i indicates the day within each period; a and b are crop-specific parameters that determine the strength of these adjustments. At the end

of yield formation, the final HI is given by:

$$HI = f_{ante}\ f_{post}\ HI_{adj} \tag{2.22}$$

Fertility Stress

Unlike crop models such as DSSAT (Jones et al., 2003) and APSIM (Holzworth et al., 2014; Keating et al., 2003), which contain modules to simulate nutrient balances and crop responses to fertilizer application rates, AquaCrop does not compute nutrient balances or take application rates as input (Steduto et al., 2009). Instead, AquaCrop requires observed maximum canopy cover, a qualitative description of the observed canopy decline, and relative biomass (the ratio between biomass with limited fertility and stress-free biomass) to employ a semi-quantitative approach that derives the relationship between fertility stress and relative biomass for the specific location (Van Gaelen et al., 2015). Although AquaCrop matches or outperforms some nutrient-balance-based crop models for maize with fertility stress in the evaluation analysis conducted by Van Gaelen et al. (2015), it neither accepts fertilizer application rates as input nor provides recommendations for fertilizer application. To facilitate large-scale yield simulation and comparison with other global gridded crop models, and to comply with the GGCMI Phase 1 protocol, we incorporate an equation of relative biomass as a function of total fertilizer application rates (sum of nitrogen, phosphorus, and potassium nutrients) to represent the crop response to fertility stress. Assuming that crop growth responds to fertilizer application with an exponential growth followed by a plateau (Cerrato & Blackmer, 1990), we apply the Mitscherlich equation (Balba & Bray, 1956; Dhanoa et al., 2022), which is developed to describe crop yield response to nutrient application, to link the GGCMI Phase 1 maize fertilizer rates with the ratio of biomass simulated under perfect irrigation by pAPSIM between *fullharm* (harmonized fertilized inputs) and *harmnon* (no nutrient limitation) experiments. pAPSIM is selected because it demonstrates a high correlation coefficient and low bias when simulating U.S. maize yields, and exhibits a high R^2 value of 0.76 for the relationship between fertilizer application rate and relative biomass. The derived equation relating fertilizer application rate to relative biomass is:

$$B_{rel} = 1 - 0.92\ * \exp(0.0069\ *\ F\ -\ 0.38) \tag{2.23}$$

where B_{rel} represents the relative biomass ranging from 0 to 1, which is the ratio of biomass under nutrient-limited conditions to nutrient stress-free biomass, and F is the fertilizer application rate in kilograms per hectare per year (kg ha^{-1} yr^{-1}).

Field-scale Calibration

The default field-scale parameterization for maize in AquaCrop is determined from a six-year field experiment dataset and a four-year canopy dataset collected at the University of California, Davis (Hsiao et al., 2009). However, the experimental site at Davis is characterized by a Mediterranean climate and relatively homogeneous soil conditions (ranging in texture from loam to silt loam). As a result, the initial parameter set may not adequately capture the wide diversity of climate, soil types, and other environmental conditions encountered across broader spatial scales.

2.2.2 Simulations Description

We run simulations at the county level to calibrate and evaluate the model against county-level yield data. We select 17 counties to conduct a global sensitivity analysis to identify the model parameters to which the output is most sensitive and that must be carefully calibrated. These 17 counties are also used for calibration (more details are provided below). We then compare the field-scale calibration version of AquaCrop and the new county-level calibration version over the entire U.S., with the model run under rainfed conditions in counties with historical rainfed maize production and under perfect irrigation in counties with historical irrigated maize production. To simulate maize growth under perfect irrigation conditions, water is added to the soil reservoir to ensure soil water content in each compartment reaches its field capacity.

We also run gridded simulations of the county-level calibrated model to evaluate it against 14 global gridded crop models (GGCMs) under the GGCMI Phase 1 protocol (Müller et al., 2019). The simulation protocol identifies three harmonization scenarios: *default*, *fullharm*, and *harmnon*, which represent different harmonization levels. The *default* setup allows modeling groups to run simulations for the default setup of their model. The *fullharm* configuration requires modeling groups to apply harmonized grid-cell specific and crop-specific growing seasons and fertilization inputs. The *harmnon* configuration also requires harmonized growing seasons but assumes unlimited nutrient supply. However, soil property inputs are not harmonized in GGCMI Phase 1 since GGCMs

have very diverse requirements for soil variables. Similarly, the calibration method, target datasets, temporal and spatial scales of calibration, and calibrated parameters are not consistent across GGCMs. Site-based crop models are usually calibrated at field scale, while ecosystem models are either uncalibrated or calibrated at large spatial scales, i.e., global or national (Müller et al., 2019).

2.2.3 Data Description

Yield and Irrigation Data

We use annual county-level maize yield data from the United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS, USDA NASS, 2023) to calibrate AquaCrop and conduct a county-level evaluation. For some counties, yield observations are explicitly reported as irrigated or non-irrigated, whereas most counties do not specify the irrigation condition of the yields. When irrigation information is not available, we rely on the high-resolution gridded (5 arc min) irrigated and rainfed harvested areas for maize from the MIRCA2000 dataset (F. T. Portmann et al., 2010), which we regrid to the county scale.

When comparing AquaCrop gridded simulations with the GGCMI Phase 1 dataset, we use U.S. national maize yields published by the FAO for national-level analysis and state-level yield records provided by USDA NASS for state-level analysis. We also use the gridded irrigated and rainfed harvested areas for maize from the MIRCA2000 dataset, consistent with GGCMI Phase 1.

We compute the area-weighted yield for each grid cell or county with the following equation:

$$Yield_{i,t} = \frac{Yield_{i,ir,t} * area_{i,ir} + Yield_{i,rf,t} * area_{i,rf}}{area_{i,ir} + area_{i,rf}} \tag{2.24}$$

where subscript i refers to the index of the grid cell or county within the contiguous U.S., and t refers to the year. The subscripts ir and rf represent irrigated and rainfed, respectively.

Climate Data

Climate inputs required by AquaCrop include daily maximum temperature, minimum temperature, precipitation, and reference evapotranspiration (ETo). In this study, we use the high-resolution gridded meteorological dataset gridMET (Abatzoglou, 2013) and spatially average its climate variables over each county to generate county-level climate inputs. These inputs are used to conduct

the global sensitivity analysis, calibrate the model, and evaluate its performance against county-level yield data. The gridMET dataset is derived from the PRISM observational climate data (Daly et al., 2008) with desirable temporal attributes from the NLDAS-2 regional reanalysis (Mitchell et al., 2004), and contains daily climate variables at a resolution of ∼4 km over the contiguous United States.

To facilitate the comparison between AquaCrop and other gridded crop models, we follow the GGCMI Phase 1 protocol and use the AgMERRA climate forcing dataset (Ruane et al., 2015) to produce gridded yield simulations. AgMERRA combines NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA, Rienecker et al., 2011) with multiple gridded and in situ observational datasets to produce daily climate variables with a spatial resolution of 0.25 degrees at the global scale spanning 1980 to 2010. With its improved representation of daily precipitation distributions and extreme events, and calibration using station data from agriculturally important regions, AgMERRA is particularly well suited for applications in agricultural modeling (Ruane et al., 2021).

Soil Data

AquaCrop identifies the soil type based on the percentages of clay, sand, and silt. For the purpose of the global sensitivity analysis, model calibration, and model evaluation, we rely on soil property data from the Gridded Soil Survey Geographic Database (gSSURGO, Soil Survey Staff, 2024) and aggregate them for each county. gSSURGO merges soil geographical data from the National Cooperative Soil Survey (NCSS) and soil attributes from the National Soil Information System (NASIS), providing gridded soil properties and characteristics at a resolution of 30 m for the conterminous United States.

Soil input is not completely harmonized across GGCMs in GGCMI Phase 1. Many GGCMs employ one of two widely used global gridded soil datasets: the International Soil Reference and Information Centre–World Inventory of Soil Emission Potentials (ISRIC-WISE) soil profile dataset (Batjes, 2006), or the Harmonized World Soil Database (HWDS Fischer et al., 2008). We choose to use HWSD (Wieder, 2014) for its ease of implementation in AquaCrop.

Fertilizer Application Data

In both county-level and gridded simulations, fertilizer application follows the GGCMI Phase 1 fertilizer input, which consists of combined published data on application rates of mineral fertilizers (Mueller et al., 2012) and manure (Foley et al., 2011; Potter et al., 2010) to produce a global gridded dataset at $0.5^\circ \times 0.5^\circ$ of crop-specific fertilizer rates of nitrogen, phosphorus (reported as P_2O_5), and potassium (reported as K_2O) nutrients (Elliott et al., 2015). In the *harmnon* experiment, we assume perfect fertilization and set the relative biomass B_{rel} to 1, thereby removing any stresses related to imperfect fertilization.

Crop Calendar

Consistent with GGCMI Phase 1, we use planting and harvesting dates from the global crop calendar dataset developed by Sacks et al. (2010), which compiles observations from different sources in the 1990s and early 2000s, presenting typical planting and harvesting dates at the sub-national level.

2.2.4 Sensitivity Analysis and County-level Calibration

To calibrate the model efficiently, we first identify the parameters to which the model output is most sensitive by performing a global sensitivity analysis. To perform the global sensitivity analysis, and then the county-level calibration, we select counties across the U.S. where maize has historically been grown and that display diverse climatic conditions and irrigation management (i.e., rainfed maize production and irrigated maize production).

To select these counties, we classify all maize-growing counties that have observed yield data for at least 20 consecutive years during 1980–2015 into three types: irrigated, rainfed, and mixed. We start by examining USDA NASS maize statistics, which report irrigated yields and rainfed yields for some counties. We classify as irrigated the counties where only irrigated yields are reported, as rainfed the counties where only rainfed yields are reported, and as mixed the counties where both irrigated and rainfed yields are reported. For counties without any specification on the irrigation system, we aggregate the gridded monthly irrigated and rainfed maize harvested areas from the MIRCA2000 dataset over each county. We classify as irrigated the counties where more than 90% of the harvested area is irrigated, as rainfed the counties where more than 90% of the harvested

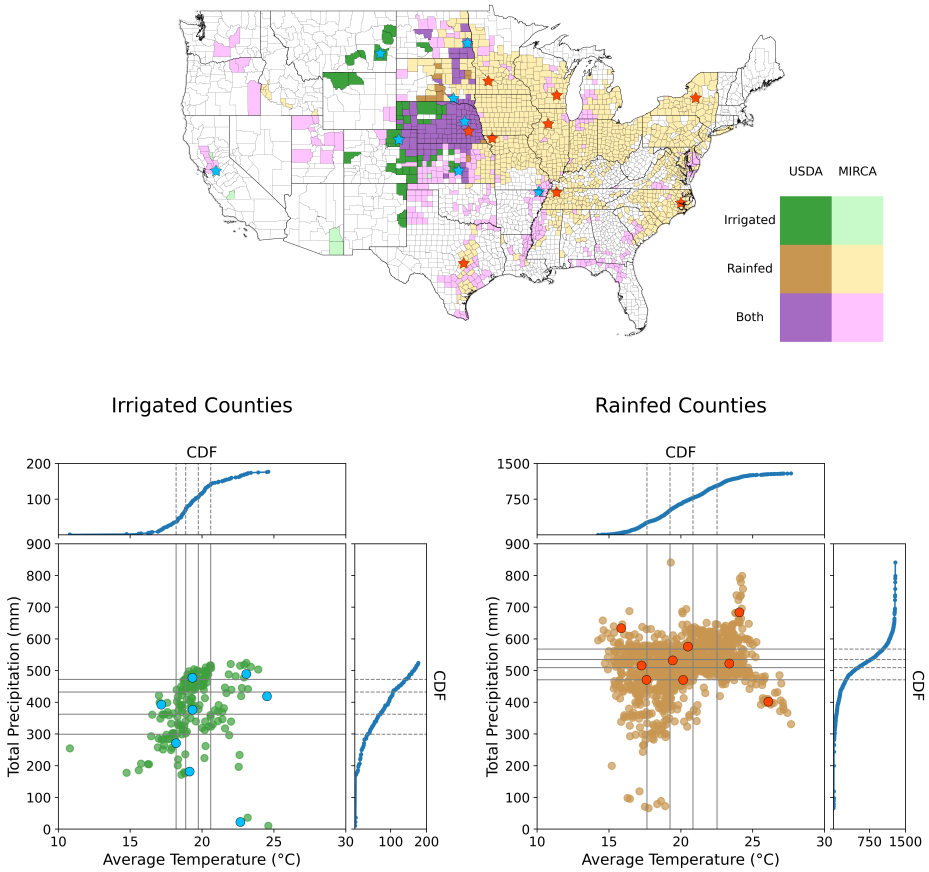

Figure 2.1: U.S. maps of maize-producing counties that have USDA yield records of at least 20 consecutive years during the period of 1985-2015 and temperature-precipitation plots over maize growing season for rainfed (**green**) and irrigated (**yellow**) counties. The types (i.e., rainfed or irrigated) of counties with light colors in the map are identified according to irrigated and rainfed crop areas provided by the MIRCA2000 dataset. If the ratios of irrigated areas or rainfed areas to the total cropland area do not exceed 90%, it is colored light **purple**. Types of counties with dark colors are identified based on the irrigated and non-irrigated yields reported by USDA. If both irrigated and non-irrigated yields were reported for a county, it is colored dark **purple**. The solid lines in temperature-precipitation plots partition counties' total precipitation and average temperature by 20%, 40%, 60%, and 80% percentiles. The top and right side plots are the cumulative distribution function plots of climate variables versus number of counties.

area is rainfed, and as mixed the remaining counties. When selecting counties with irrigated maize, we select from irrigated and mixed counties; when selecting counties with rainfed maize, we select from rainfed and mixed counties (see Fig. 1).

To represent diverse climatic conditions, we partition for each county the average precipitation and average temperature over the growing period by percentiles of 20%, 40%, 60%, and 80%, and assign five classes: high (H), medium high (MH), medium (M), medium low (ML), and low (L). This generates 25 bivariate precipitation–temperature climate categories (e.g., H–M refers to high precipitation and medium temperature). For rainfed maize and for irrigated maize, we select nine

counties, each corresponding to one of the following categories: H–H, H–M, H–L, M–H, M–M, M–L, L–H, L–M, L–L. Due to the absence of irrigated counties in the H–L category (i.e., high precipitation and low temperature), only eight counties are selected for the sensitivity analysis and calibration of irrigated maize. Within each representative climate category, we choose the county with relatively high historical maize production to ensure that the sensitivity analysis and calibration focus on agriculturally significant areas. To capture physiographic variation and avoid spatial clustering, we also prioritize counties that are geographically distant from one another. The final 17 counties selected for the global sensitivity analysis and county-level calibration are shown in Fig. 1.

To conduct the global sensitivity analysis, we use Sobol's method (Sobol, 2001), a variance-based global sensitivity analysis, to calculate the first-order and total-order sensitivity indices for each input parameter. The first-order index represents the individual relative contribution of a parameter to the total variance of the model output, and the total-order index accounts for both the individual contribution of the parameter and its interactions with other parameters. The threshold of Sobol's sensitivity indices to identify sensitive parameters is not clearly defined, although previous work often uses 0.05 (X.-Y. Zhang et al., 2015). In our study, we include parameters with total sensitivity indices above 0.01 to expand the dimensionality of the calibration space (Fig. 2).

We apply the Shuffled Complex Evolution algorithm developed at the University of Arizona (SCE-UA, Duan et al., 1993, 1994) to calibrate and optimize AquaCrop. This calibration method randomly samples parameter sets from defined parameter ranges and partitions them into complexes to broadly explore the parameter space. Parameter sets are updated by nudging toward the one with the best performance or by taking random jumps to explore new parameter regions. Updated parameter sets are merged and re-partitioned into complexes repeatedly until convergence is achieved. The SCE-UA approach balances exploration and exploitation when searching for the best-performing parameter set, and its use of complexes allows us to calibrate AquaCrop in parallel. The calibration ranges are defined based on the AquaCrop reference manual, previous calibration studies for AquaCrop, and maize phenology. For parameters with ranges specified for maize in the reference manual, such as *GDD_sene* and *Water_Productivity*, their ranges are directly adopted in the calibration (Raes et al., 2023a). Parameters representing soil water stress thresholds, such as *P_up_exp*, are inherently bounded between 0 and 1, and their full feasible ranges are used. The lethal maximum temperature reported in a review study that summarizes critical thresholds for

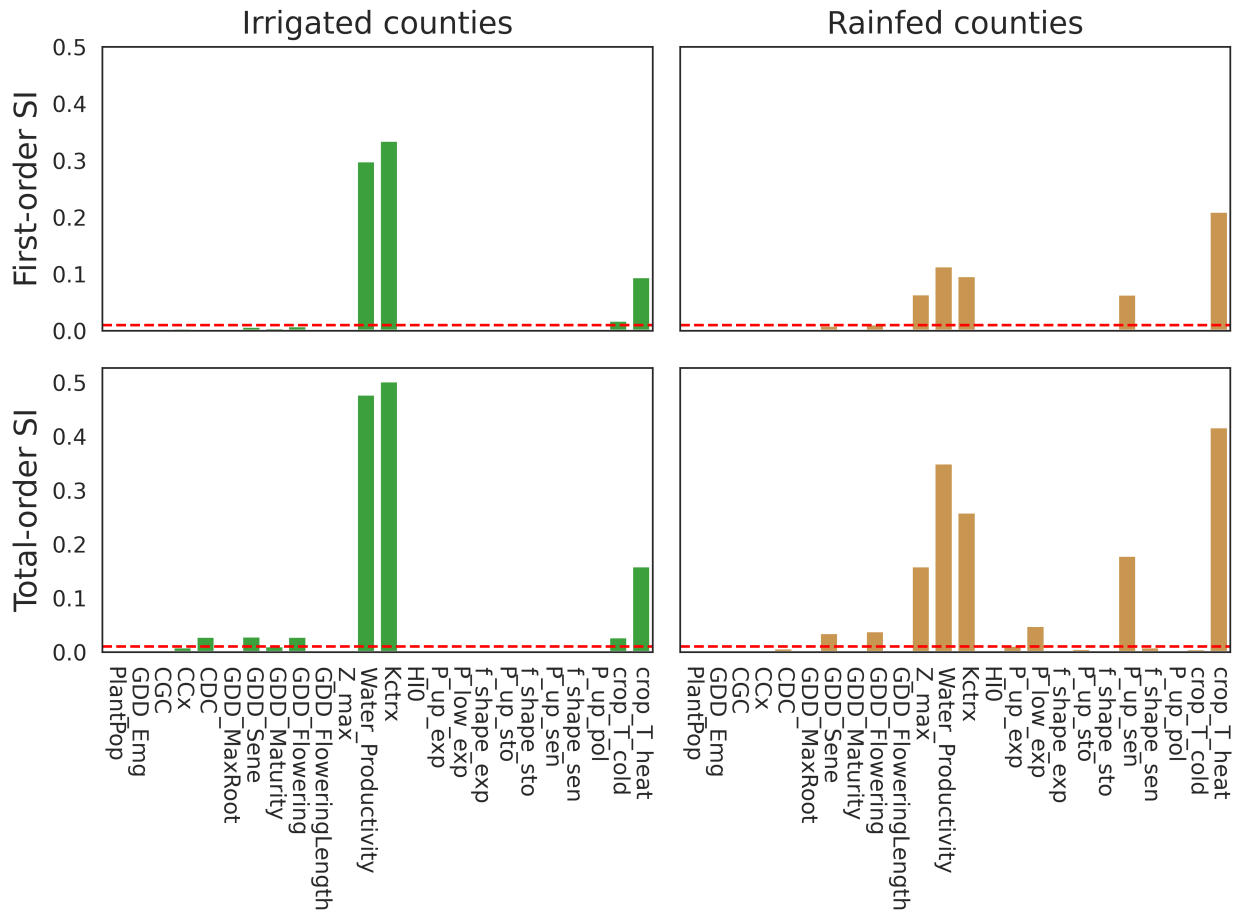

Figure 2.2: Sobol's sensitivity indices (first-order and total-order) for irrigated counties (**green**) and rainfed counties (**yellow**) of parameters.

maize and rice (Sánchez et al., 2014) is used as the upper bound for *Crop_T_heat*, and we set the default minimum temperature for cold stress as the lower bound for *Crop_T_heat* to allow broader exploration of the parameter space. We show the parameters values resulting from the county-level calibration, along with the default field-scale calibration values in Table 1.

2.3 Results

2.3.1 County-level Performance

To evaluate the performance of AquaCrop using both field-scale calibrated parameters and county-level calibrated parameters against USDA county-level maize yields, we run simulations for both irrigated and rainfed counties that have at least 20 consecutive years of data during 1985–2015. All county-level simulations use the gridMET climate dataset (Abatzoglou, 2013), the gSSURGO

Table 2.1: Calibrated and default values of calibrated AquaCrop parameters

Parameter	Description	Default	Calibrated
CDC	Canopy decline coefficient (fraction per growing degree day)	0.01	0.013
GDD_Sene	GDD from sowing to start of senescence	1400	1294
GDD_Flowering	GDD from sowing to flowering	880	624
Z_max	Maximum root depth (m)	2.3	1.82
Water_Productivity	Water productivity normalized for ETo and CO_2 (g/m^2)	33.7	31.1
Kctrx	Crop coefficient when canopy is complete but prior to senescence	1.05	1.01
P_up_exp	Soil water depletion threshold for canopy expansion – Upper threshold	0.14	0.48
P_low_exp	Soil water depletion threshold for canopy expansion – Lower threshold	0.72	0.92
P_up_sto	Soil water depletion threshold for stomatal control – Upper threshold	0.69	0.27
P_up_sen	Soil water depletion threshold for canopy senescence – Upper threshold	0.69	0.95
Crop_T_heat	Max air temperature above which pollination fails (heat stress) (°C)	40	43

soil dataset (Soil Survey Staff, 2024), and the fertilization application rates provided by GGCMI Phase 1 (Müller et al., 2019) as inputs. To obtain county-level yields for all counties meeting this criterion, we follow the GGCMI Phase 1 approach: running county-level simulations under both perfectly irrigated and purely rainfed conditions, and then calculating the area-weighted yield based on irrigated and rainfed crop areas provided by MIRCA2000.

Compared with USDA county-level maize yield records, the field-scale calibrated AquaCrop exhibits a systematic spatial trend of substantial underestimation in several northern states and overestimation in the southeast, with a U.S. production-weighted mean absolute bias of 28%(Fig. 3a). The county-level calibrated AquaCrop substantially reduces the absolute bias percentage (Fig. 3b) to 16%. This indicates that county-level calibrated parameters account more effectively for spatial heterogeneity in environmental and agronomic conditions and significantly mitigate systematic error in AquaCrop. In addition, the county-level calibrated version demonstrates improved performance in the Corn Belt region. However, the spatial pattern of underestimation in the northwestern Corn Belt and overestimation in the southeast is still present in county-level calibrated version, although much more dampened.

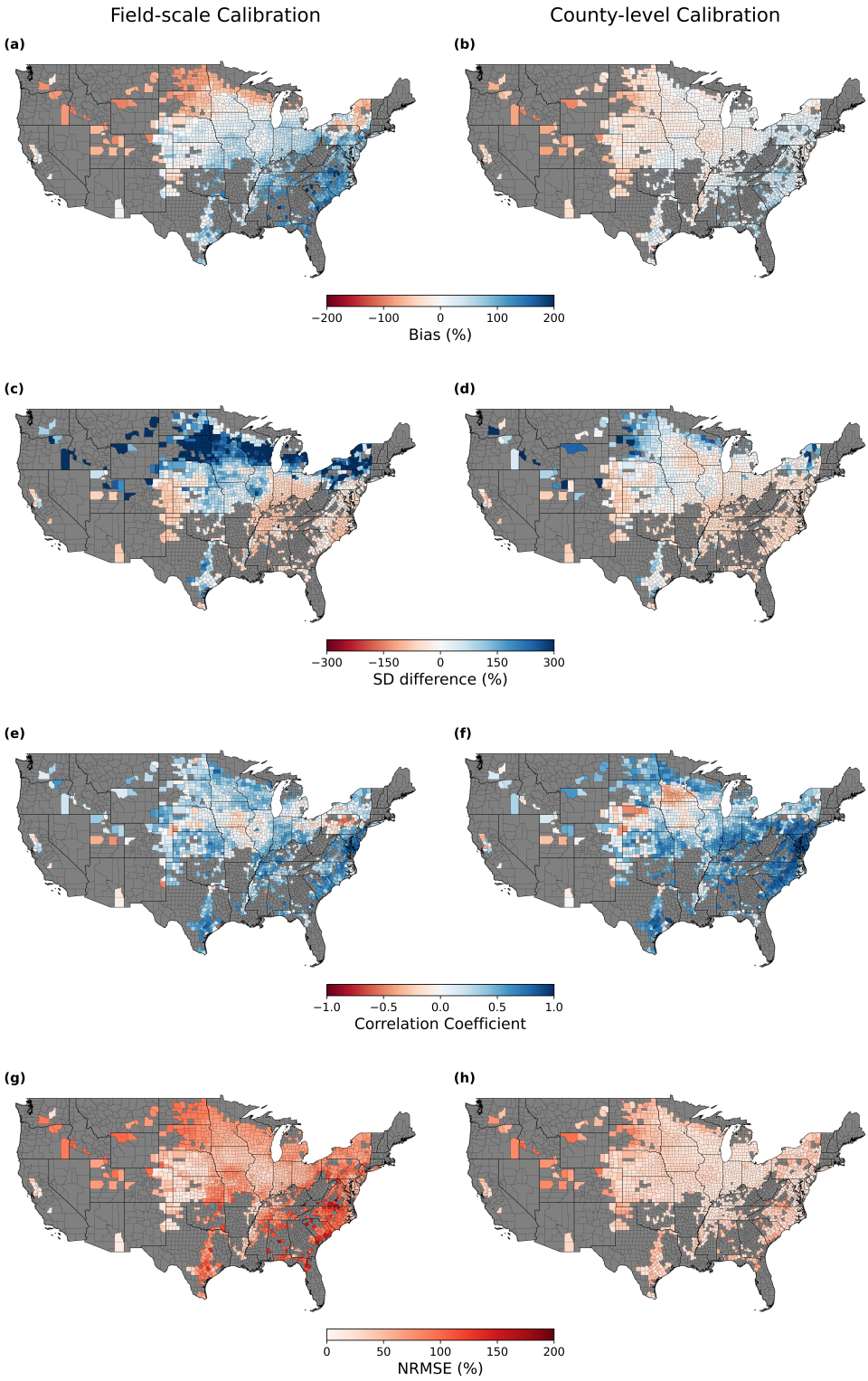

Figure 2.3: U.S. maps of county-level evaluation metrics for all counties that have USDA yield records of at least 20 consecutive years during the period of 1985-2015. **a**, **c**, **e**, and **g** show county-level bias in percentage, standard deviation difference in percentage, correlation coefficient, and root mean square error normalized by averaged observed yield for field-scale calibrated simulations. **b**, **d**, **f**, and **h** show county-level evaluation metrics for county-level calibrated simulation.

To assess AquaCrop's performance in capturing the temporal variability of maize yields, we compute the percentage difference between the simulated and observed standard deviations (SD) at the county level (Fig. 3c and 3d). The field-scale calibrated AquaCrop demonstrates strong spatial heterogeneity across states, with substantial overestimation of yield variability in some northern states such as South Dakota, Wisconsin, Michigan, and New York (Fig. 3c). After county-level calibration, the magnitudes of SD differences are reduced across nearly all counties, with particularly large improvements in those northern states. The U.S. production-weighted average absolute relative difference in county-level SD decreases from 127% to 36%. Similar to the analysis of bias, the calibrated AquaCrop performs well in the core Corn Belt region, where SD differences after calibration are clearly smaller than in other areas. However, calibration also leads to an overall underestimation of variability. The number of counties with negative SD differences increases from 700 to 1104. This shift from overestimation to underestimation suggests that the initial parameterization may have exaggerated maize's sensitivity to interannual climate variability, while the county-level calibrated parameters may reduce this sensitivity too much. Despite this limitation, the overall reduction in SD differences and the improved agreement in the Corn Belt highlight the model's improved ability to reproduce yield variability.

The maps of county-level correlation coefficients (Fig. 3e and 3f) indicate the model's ability to capture temporal trends in yields. Simulated yields from the field-scale calibrated AquaCrop have weak correlation with observed yields, particularly in counties in Iowa, Nebraska, Missouri, Pennsylvania, and Ohio (Fig. 3e). After calibration, the temporal correlation improves across most counties, especially in Missouri, Illinois, Pennsylvania, and the east coastal regions (Fig. 3f), with the U.S. average correlation coefficient increasing from 0.32 to 0.42. The number of counties with correlation coefficients above 0.5 nearly doubles from 320 to 611, while the number of counties with negative correlation drops from 247 to 220. Spatially, the strongest improvements are concentrated in the Great Lakes region and the east coast, while counties in Nebraska and Iowa show only modest gains or even deterioration.

The county-level RMSE normalized by observed mean yields (NRMSE, in percentage) (Fig. 3g and 3h) indicates the overall accuracy of AquaCrop in simulating maize yields. Across nearly all simulated counties (except for some in Colorado, Nebraska, and Kansas) the field-scale calibrated AquaCrop exhibits high NRMSE, with a U.S. production-weighted average of 47%, indicating

considerable deviation from observed yields and highlighting poor model accuracy (Fig. 3g). The county-level calibration successfully reduces the NRMSE to 26%, although counties in the West show limited improvements, suggesting that including more counties from these areas will be necessary in future calibrations.

In terms of regional variation in model performance, our analysis reveals differences in the sources of deviation between observed and simulated yields after county-level calibration. Along the east coast, AquaCrop simulations show strong correlation but large bias, implying that the model can capture overall yield fluctuations but systematically overestimates or underestimates yield levels. However, despite the strong correlation, this region exhibits relatively large SD differences, indicating that AquaCrop does not always capture the magnitude of interannual variability. Reliable performance is observed in counties with low bias, strong correlation, and similar SD. For example, in Indiana, the county-level calibrated AquaCrop successfully reproduces both the magnitude and the interannual variability of yields. In contrast, counties with large bias, large SD differences, and low or negative correlation are found in Wyoming, Idaho, Oregon, and Utah, indicating consistent yield misestimation and failure to reproduce temporal dynamics in these regions. This spatial pattern provides critical insight into the regional strengths and limitations of AquaCrop and highlights areas where further refinement or region-specific calibration may be needed.

We also conduct separate analyses for rainfed and irrigated counties to validate AquaCrop's performance under purely rainfed and perfectly irrigated conditions. As described in the methods section, counties with USDA-specified irrigated or non-irrigated yields are included in the irrigated and rainfed analyses, respectively, even when their irrigation percentages fall between 10% and 90%. For rainfed counties, evaluation metrics demonstrate substantial improvement after county-level calibration, except in the border region shared by Nebraska, Colorado, and Kansas (Supplement Fig. S1). In this region, counties typically have high irrigation percentages, so their simulated yields are evaluated against USDA-specified non-irrigated yields for the rainfed analysis. Considering AquaCrop's robust performance in other rainfed counties, these discrepancies may reflect data classification issues rather than model limitations. For irrigated countie (mostly located in Nebraska, Colorado, and Kansas) calibration reduces bias but shows little improvement in correlation or NRMSE, highlighting the challenge of simulating irrigated maize yields (Supplement Fig. S2).

The combined reduction in yield bias, NRMSE, and SD differences, along with the improved

correlation coefficients across most simulated counties, clearly demonstrate that our calibration approach, using only 9 rainfed counties and 8 irrigated counties, remarkably improves AquaCrop's ability to reproduce both yield levels and interannual yield variability. These improvements underscore the effectiveness of county-level calibrated parameters in accounting for spatial heterogeneity in soil, climate, and management practices that are not captured by the initial parameterization. By correcting systematic spatial errors and enhancing the model's ability to reproduce interannual yield fluctuations, the county-level calibrated AquaCrop provides a more robust and reliable tool for simulating maize yield dynamics at large spatial scales.

2.3.2 Comparison with GGCMI Phase 1

To evaluate the performance of AquaCrop relative to other gridded crop models, we compare the gridded maize yield simulations from the county-level calibrated AquaCrop with the GGCMI Phase 1 dataset for the contiguous U.S., following its evaluation approach (Müller et al., 2017, 2019). The GGCMI Phase 1 dataset provides an ensemble of global gridded yield simulations from 14 crop models, offering a robust benchmark for assessing crop model performance. Through this comparison, our aim is to identify the strengths and limitations of AquaCrop at different spatial scales. Following the simulation protocol of GGCMI Phase 1, we conduct gridded maize yield simulations under both purely rainfed and perfectly irrigated conditions spanning 1980 to 2009. We use AgMERRA as climate input and HWSD as soil input, the same datasets used in GGCMI Phase 1 simulations. Rainfed and irrigated yields of each grid cell are weighted by their corresponding areas from the MIRCA2000 dataset to obtain the final maize yield for that grid cell. We provide two harmonization scenarios: *fullharm* and *harmnon*.

State-level Analysis

When comparing area-weighted state-level yields against USDA observations, we find significant spatial variation in model performance. Under the *fullharm* harmonization setup, most crop models exhibit a consistent bias pattern: underestimating yields in western states while overestimating them in eastern states (Fig. 4). For example, all models consistently overestimate maize yields in Florida while significantly underestimating yields in Idaho, which may result from inaccuracies in harvested area estimates (i.e., MIRCA2000). On the other hand, models vary considerably in

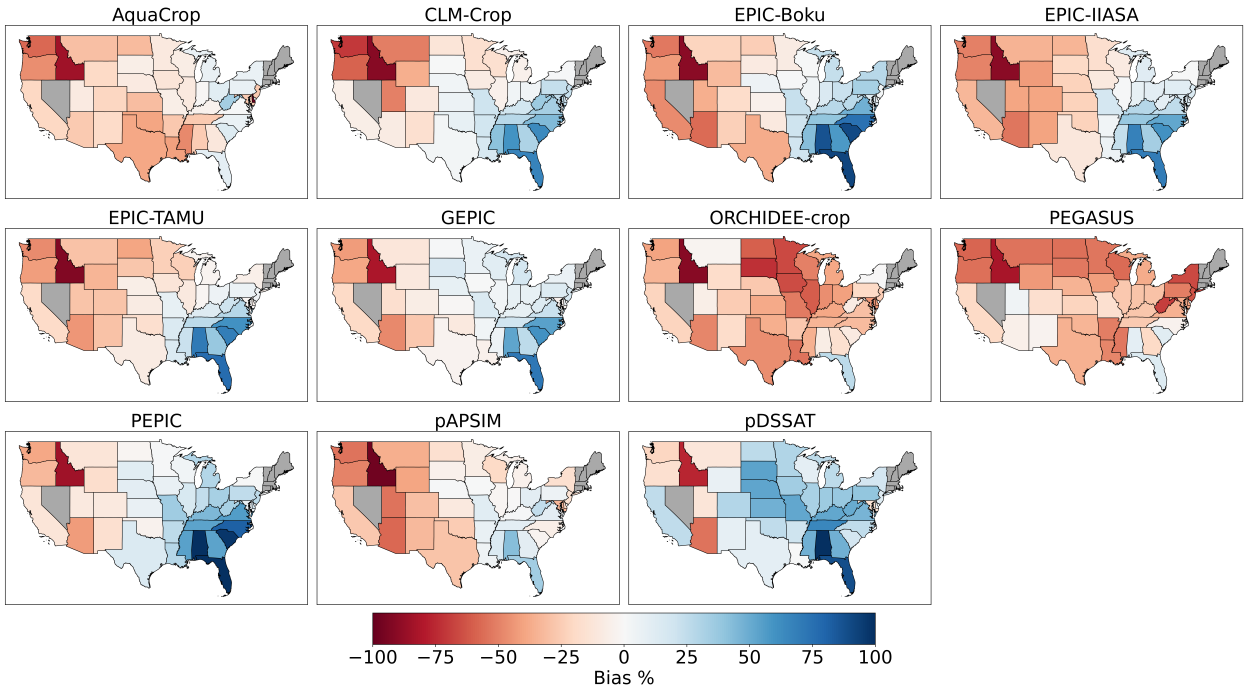

Figure 2.4: U.S. maps of state-level maize yield bias (in percentage) between the simulated yield under the *fullharm* harmonization scenario and the observation for the period of 1980-2009. GCMs that didn't provide *fullharm* simulations were not included.

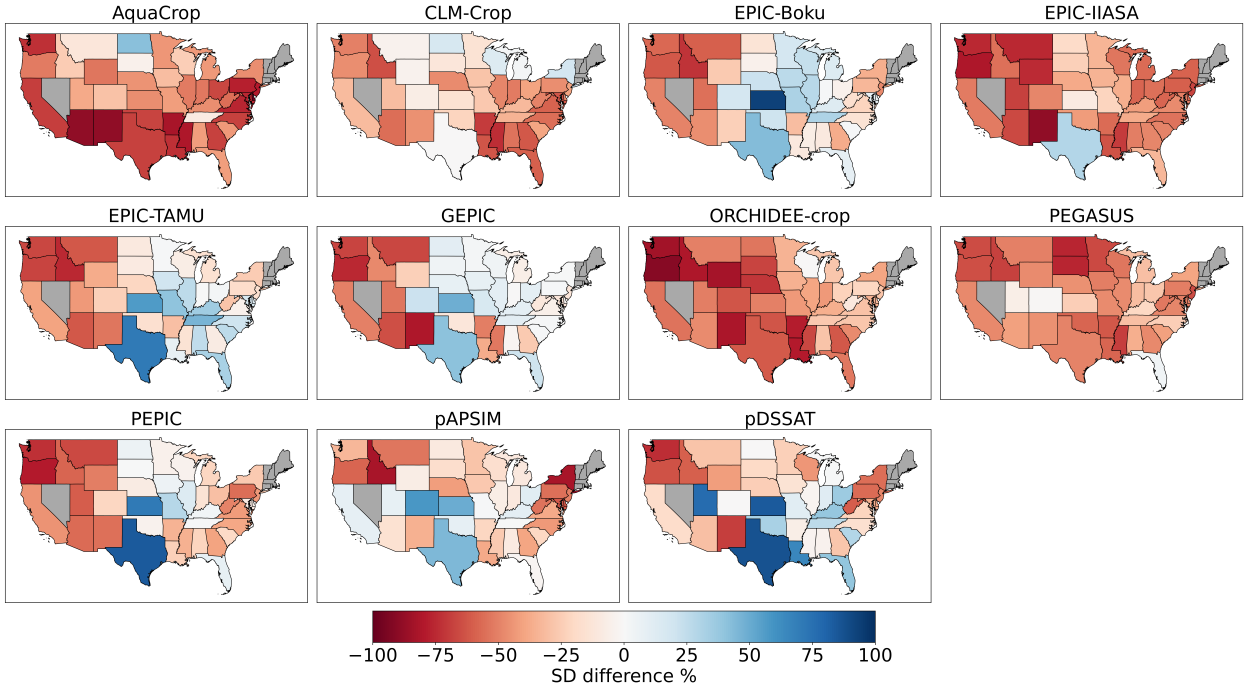

Figure 2.5: Same as Fig. 4 but showing the standard deviation difference (in percentage).

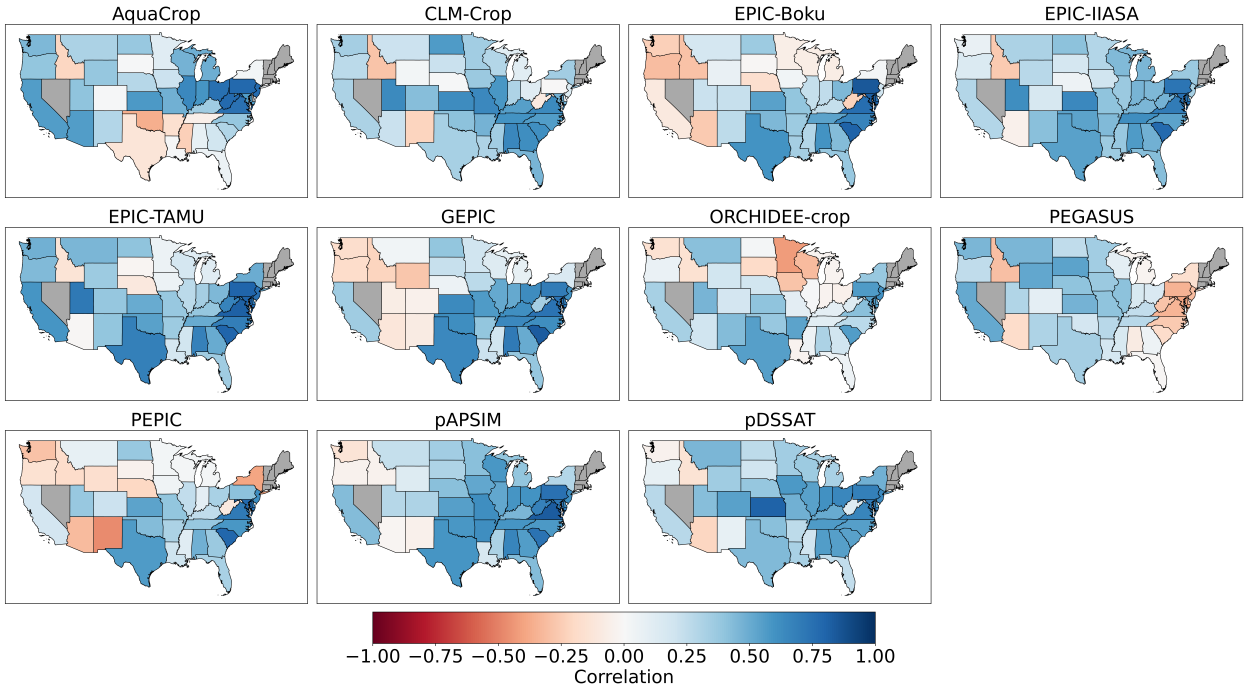

Figure 2.6: Same as Fig. 4 but showing the temporal correlation coefficient.

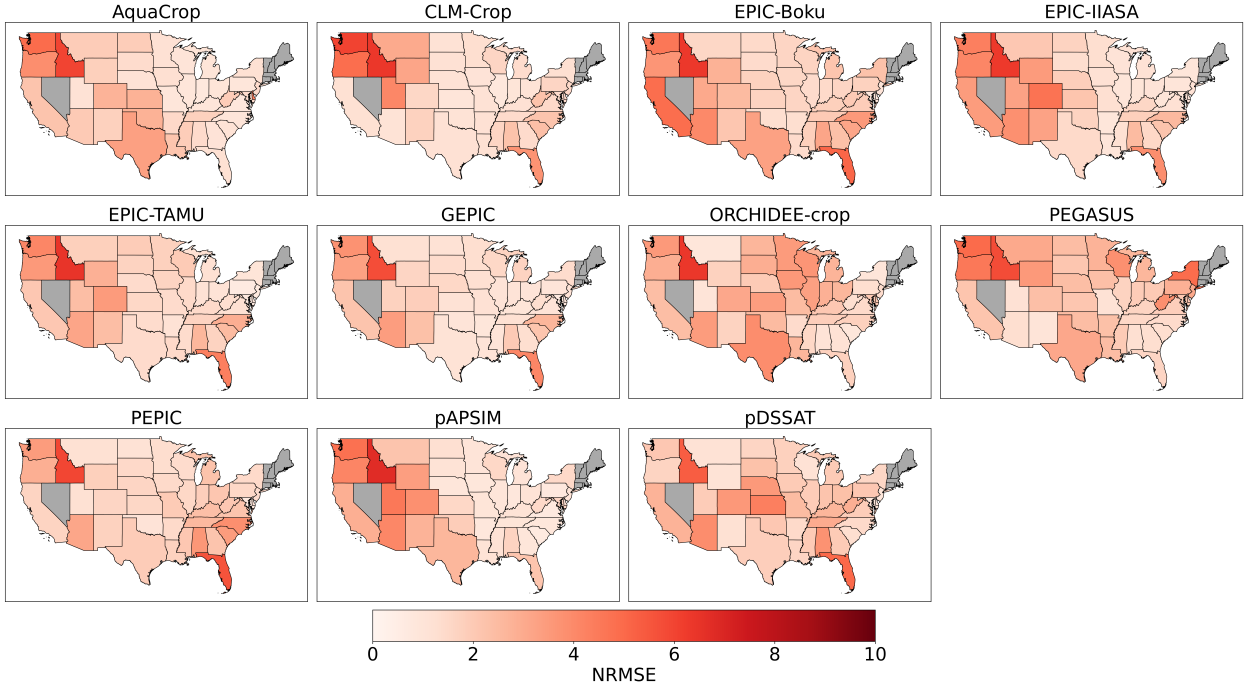

Figure 2.7: Same as Fig. 4 but showing the RMSE normalized by observed state-level standard deviation.

their ability to capture state-level yield variability (Fig. 5). AquaCrop tends to underestimate temporal variability across states, a pattern shared by ORCHIDEE-Crop, PEGASUS, CLM-Crop, and EPIC-IIASA. Conversely, models like EPIC-Boku overestimate standard deviation in the Corn Belt region. Despite differences in skill in capturing variability, most models show strong performance in reproducing temporal variations of state-level yields (Fig. 6). However, GEPIC exhibits poor temporal correlations in western states, while PEGASUS performs poorly along the east coast. In terms of overall accuracy, Idaho shows the largest deviations across all models, while most models exhibit relatively smaller RMSE normalized by standard deviation in the Corn Belt region (Fig. 7). AquaCrop demonstrates consistent performance across scenarios (*fullharm*, and *harmnon*), whereas other models show substantial variation between scenarios. For instance, PEGASUS shows opposite patterns between *fullharm* and *harmnon*: underestimating yields and standard deviation in *fullharm*, while overestimating both in *harmnon*, likely reflecting its high sensitivity to fertility variations.

National-level Analysis

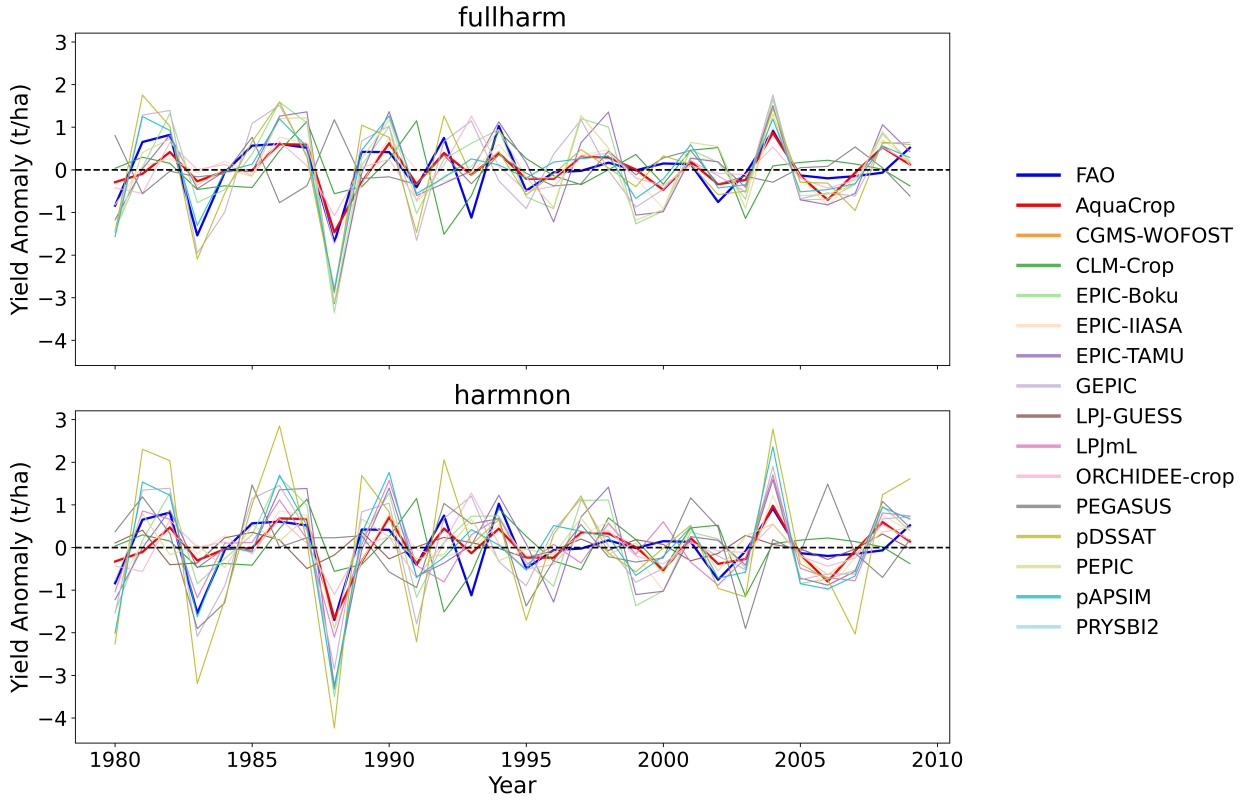

Figure 2.8: Time-series of U.S. national yield anomalies from calibrated AquaCrop simulations, 14 GGCMs, and FAO data for 3 harmonization scenarios: *default*, *fullharm*, and *harmnon*. The yield anomalies were obtained by detrending yields with 5-year moving average.

To evaluate performance at the national level, we aggregate gridded yields to national yields by weighting each grid cell by its maize production. Compared with other global gridded crop models (GGCMs) included in GGCMI Phase 1, county-level calibrated AquaCrop demonstrates competitive skill in capturing the temporal dynamics of detrended yields in both magnitude and variability across all three scenarios (Fig. 8). Under the *harmnon* setup, which assumes unlimited nutrient supply, interannual yield variability shows greater amplitude compared to the *fullharm* scenarios for both AquaCrop and most GGCMs. Without nutrient limitations, models exhibit exaggerated yield fluctuations, especially in years with extreme yields, and pDSSAT shows the most pronounced variation. Notably, in certain years such as 1993, several GGCMs exhibit yield anomalies opposite in direction to those reported in the FAO data, while AquaCrop successfully captures the anomaly.

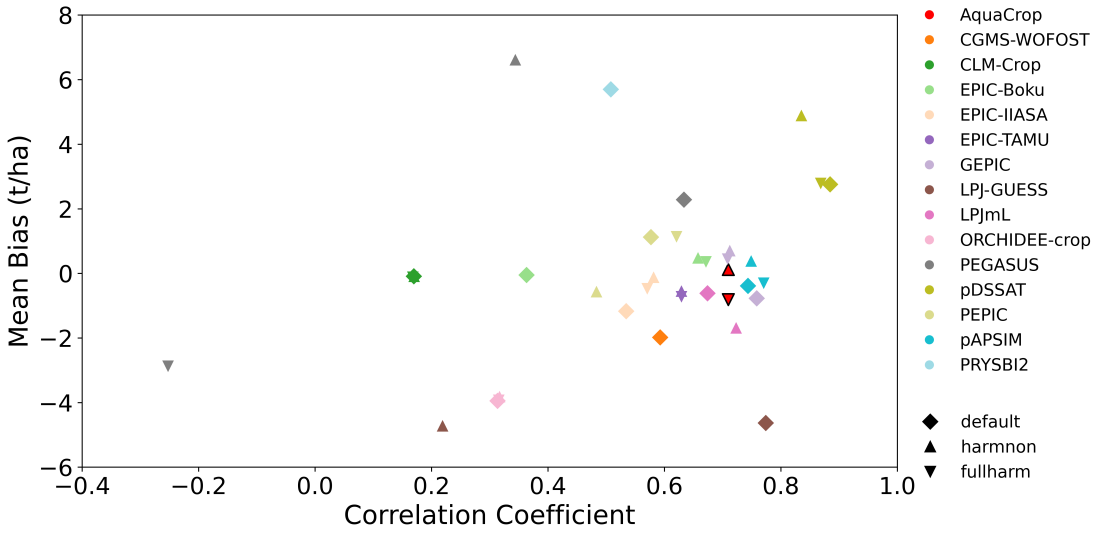

Figure 2.9: U.S. national mean yield bias and correlation coefficient for maize simulated by AquaCrop and GGCMs when evaluating against the FAO data.

The correlation coefficients of the detrended yield time series range from –0.25 (*fullharm* PEGASUS) to 0.84 (*default* pDSSAT) (Fig. 9). pDSSAT consistently outperforms the other models in reproducing maize yield trends at the national scale in this study and at the global scale in Müller et al. (2017), while PEGASUS under the *fullharm* configuration performs the worst at both scales. AquaCrop also demonstrates competitive skill in reproducing the temporal pattern of FAO data, with correlation coefficients of nearly 0.71 across the three scenarios. In terms of yield bias, which indicates a model's ability to estimate the national yield level, PEGASUS (*harmnon*) exhibits the largest overestimation (6.6 t/ha), while LPJ-GUESS shows the largest underestimation (–4.7 t/ha)

in both the *default* and *harmnon* scenarios. Although pDSSAT displays the strongest temporal correlation with the reference data, its simulations deviate substantially from the national maize yield level, with overestimation between 2.8 and 4.9 t/ha. The smallest deviation is provided by EPIC-Boku under the *default* scenario (–0.05 t/ha), though its correlation coefficient is less than 0.4. AquaCrop performs notably well with minor underestimation of U.S. baseline maize yields, particularly in the *harmnon* scenario, with a bias of only 0.12 t/ha.

Crop models show different responses to changes in harmonization settings. For models such as CLM-Crop and ORCHIDEE-Crop, biases and correlation coefficients do not change much across the three scenarios, indicating their relative insensitivity to fertilizer inputs. In contrast, simulations of PEGASUS are dramatically affected by the choice of harmonization setup. After removing nutrient limitation (i.e., *harmnon*), its correlation coefficient rises from –0.25 to 0.34, while its bias increases from –2.9 to 6.6 t/ha. Other models, including AquaCrop, show reasonable changes in bias and correlation across scenarios, indicating a balance of robustness and sensitivity to harmonization settings.

Spatiotemporal Analysis

To summarize the evaluation of AquaCrop's performance in reproducing both temporal dynamics and spatial patterns of observed maize yields, we plot the spatiotemporal-dynamic Taylor diagram using two-dimensional data (State × Year) (Fig. 10). Following the GGCMI Phase 1 evaluation approach, we detrend yield time series while preserving the mean yield for each state, and then weight them by their state-level production to construct the Taylor diagram. In terms of normalized root mean square deviation (NRMSD), the *default* GEPIC provides the best performance, and *fullharm* AquaCrop achieves the second-lowest NRMSD out of 37 model–harmonization combinations. pDSSAT demonstrates the strongest spatiotemporal correlation with state-level yield records, consistent with the national-scale analysis. However, under the *harmnon* setup, pDSSAT tends to overestimate the amplitude of yield variation, with a normalized standard deviation (NSD) of 2.36, which is also reflected in Fig. 5, resulting in a large deviation from the reference data. Overall, AquaCrop is showing high performance compared to the other GGMI models.

We also plot the temporal-dynamic-only Taylor diagram by removing each state's mean yield, and the spatial-dynamic-only Taylor diagram by averaging time series across years, to further

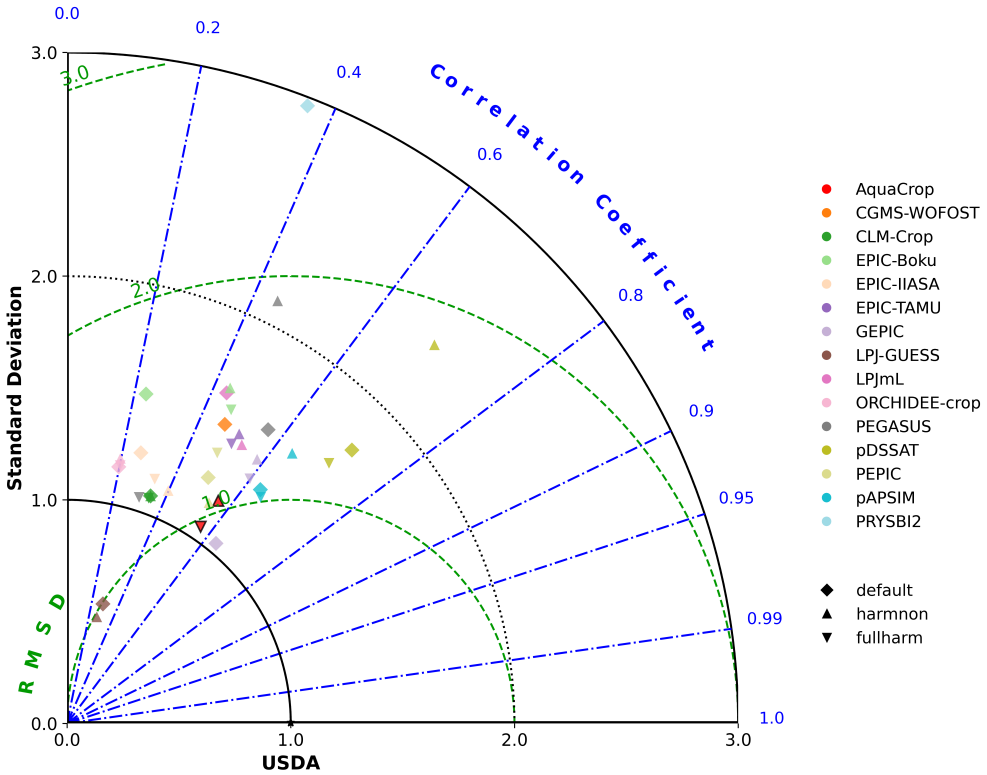

Figure 2.10: Taylor diagram of U.S. state-level yields simulated by calibrated AquaCrop and GGCMs evaluating against the USDA data records. Yields were detrended but preserving the baseline yield for each state. The standard deviation for each model was normalized by that of reference data (USDA). Individual states were weighted by their productions.

investigate the sources of yield variability (Supplement Fig. S7 and Fig. S8). We find that the spatiotemporal-dynamic Taylor diagram is similar to the spatial-dynamic-only version for most models, though with different magnitudes, suggesting that the majority of the simulation variability arises from spatial variability.

2.4 Discussion and Conclusion

Rising food demand and adverse environmental impacts on crop production continue to heighten concerns about global food security (Tilman et al., 2011). Process-based crop models such as AquaCrop, DSSAT, and APSIM are widely used to simulate crop responses to environmental conditions and management practices, assess climate impacts, and explore adaptation strategies (M. Kumar, 2016). AquaCrop, developed by the Food and Agriculture Organization, explicitly represents crop growth and responses to water stress. Despite its relatively simple structure, many studies have validated its performance (Abedinpour, 2021; Battisti et al., 2017; Feleke et al.,

2021). This balance of simplicity, robustness, and accuracy makes AquaCrop accessible to diverse users and useful for large-scale analyses. However, regional implementations often lack flexibility for modification, calibration, and integration with other datasets or models (Chen et al., 2023). Moreover, most process-based models are parameterized with data from only a few sites, raising concerns about their ability to capture the diversity of climate, soils, and management practices across large regions. Collecting in situ data at broader scales is also resource intensive. To address these limitations, we develop a gridded, Python-based implementation of AquaCrop and proposed a calibration approach using U.S. county-level yield data. We focus on maize because it is central to the human diet, animal feed, and bioethanol production (García-Lara & Serna-Saldivar, 2019), and because the United States produces nearly 30% of global maize, worth about 70 billion US dollars annually (FAO, 2024; USDA NASS, 2025).

We evaluate how a version of AquaCrop calibrated using county-level yield data only performs compared to a version calibrated at the field scale with multiple target variables. County-level calibration significantly improves performance, as the U.S. production-weighted average absolute bias decreases from 28% to 16%, the U.S. production-weighted average SD error decreases from 127% to 36%, the U.S. production-weighted average NRMSE drops from 47% to 26%, and the average temporal correlation coefficient rises from 0.32 to 0.42. Importantly, these gains are achieved with calibration based on only 9 rainfed and 8 irrigated counties, showing that carefully selected sites, along with the selection of parameters for calibration based on a global sensitivity analysis, can improve model performance across most maize-producing regions of the U.S. by better capturing environmental and management heterogeneity.

Despite the standardized protocols of GGCMI Phase 1, global gridded crop models (GGCMs) vary widely in calibration procedures (Müller et al., 2019). Ecosystem models such as LPJ-GUESS and ORCHIDEE-Crop are not calibrated; LPJmL is calibrated with FAO national yields; and others, like PEGASUS and PRYSBI2, rely on global gridded yield datasets (Iizumi et al., 2014; Monfreda et al., 2008). Site-based models often use field-scale calibrations, though some EPIC models are calibrated at national scales (Deppermann et al., 2018; Sauer et al., 2010). Temporal calibration coverage also differs: most GGCMI Phase 1 calibrations use single-year reference data around 2000, except PRYSBI2, which uses 1982–2006. These differences limit models' ability to capture interannual variability, as reflected in Müller et al. (2017). Our analysis, like theirs, shows that spatial

variability dominates overall yield variability, but we also find that temporal variability contributes more at the U.S. state level than in global analyses. This is expected, as spatial heterogeneity declines at smaller scales. Still, accurate simulation of interannual fluctuations is critical for applications such as harvest forecasting, early warning systems, and policy planning (Maestrini & Basso, 2018). Improving temporal dynamics should therefore remain a priority alongside spatial accuracy.

Our calibration approach enhances AquaCrop's regional representativeness by incorporating two decades of observed county-level yield data from diverse climatic zones and major production areas. Unlike traditional site-specific calibrations that focus on a few years of in situ data (Seidel et al., 2018), our approach captures broader spatial and temporal variability. Similar regionalized calibrations in Europe (Angulo et al., 2013) and global efforts with EPIC (Xiong et al., 2016) have shown the value of extending beyond site-level data, though they also highlight challenges of dataset requirements and computational demand (Pasquel et al., 2022). We expand the calibration parameter space more than many previous regional studies (11 parameters) based on Sobol's sensitivity analysis. Using 20 years of records from 17 counties strategically chosen to span precipitation–temperature regimes, we achieve substantial performance gains while keeping computational cost feasible. Adding more counties would likely improve results further but at greater cost. Our strategy balances accuracy and practicality, providing a pathway for robust large-scale calibration of AquaCrop.

We acknowledge several limitations. Our calibration relies on the accuracy of USDA yield records and MIRCA2000 harvested area data, and the focus on maize means further work is needed to extend this approach to other crops. Irrigation representation also remains simplified, suggesting the need for more detailed water-use datasets to improve performance in irrigated systems. In addition, our comparison to other GGCMI models is based solely on the AgMERRA climate forcing dataset and there is clear evidence of strong regional influence of the choice of the climatic forcing datasets on crop model simulations. Nonetheless, the consistent evaluation across simuations with the same climatic forcing dataset should provide a fair assessment of the strengths and weaknesses of AquaCrop compared to the other GGCMI models.

By improving both spatial representativeness and temporal fidelity, our approach enhances the utility of AquaCrop for applications ranging from regional adaptation planning to national-scale assessments of climate impacts. The calibrated framework can also support model intercomparison projects such as AgMIP and GGCMI by providing a benchmark that bridges field-scale processes

and large-scale decision needs. Future work should expand the calibration framework to additional crops, incorporate dynamic management practices such as cultivar choice and fertilizer timing, and explore integration with machine learning approaches to further improve predictive skill.

Supplement

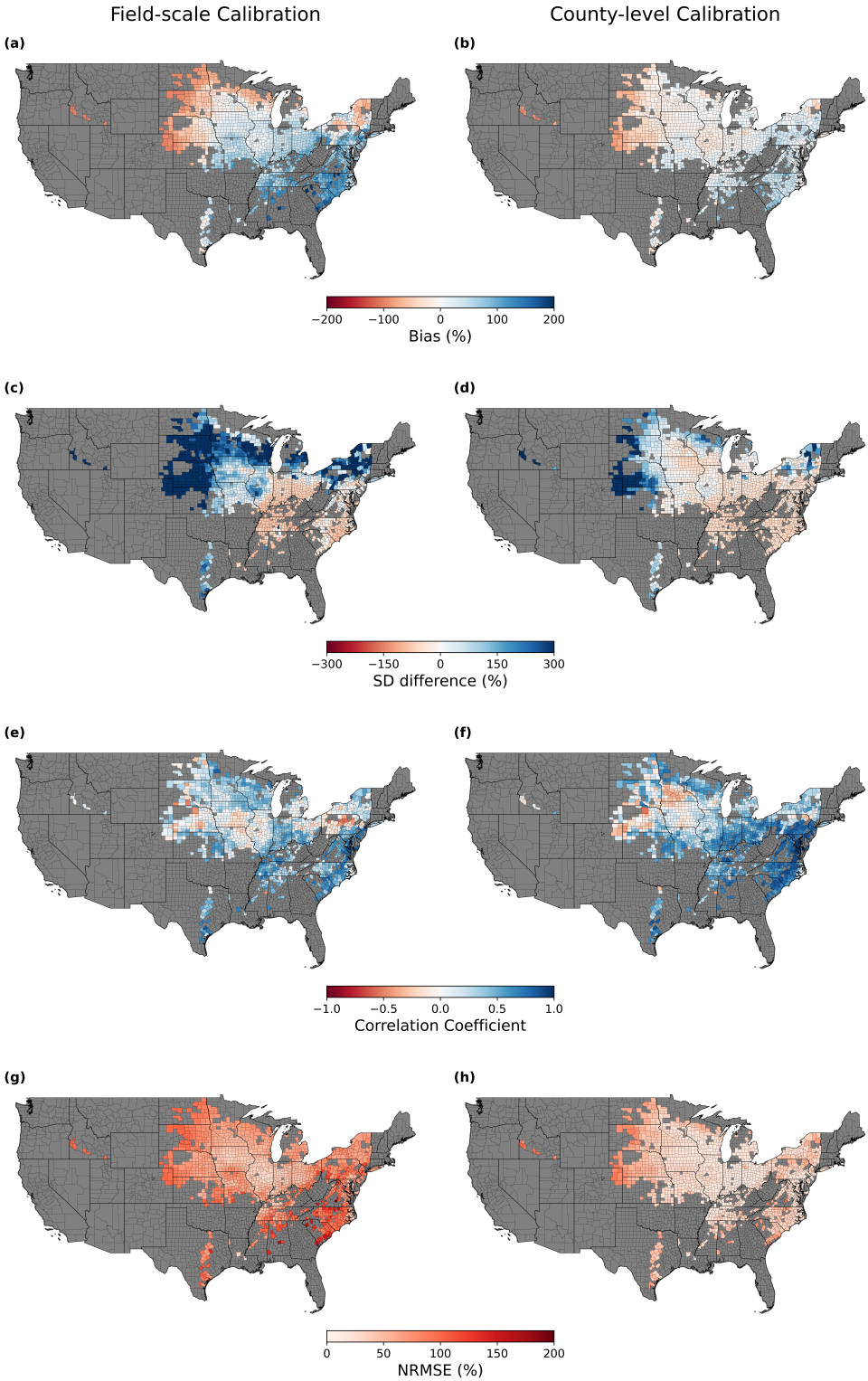

Figure S1: As Figure 2.3 in the main text, but for rainfed counties.

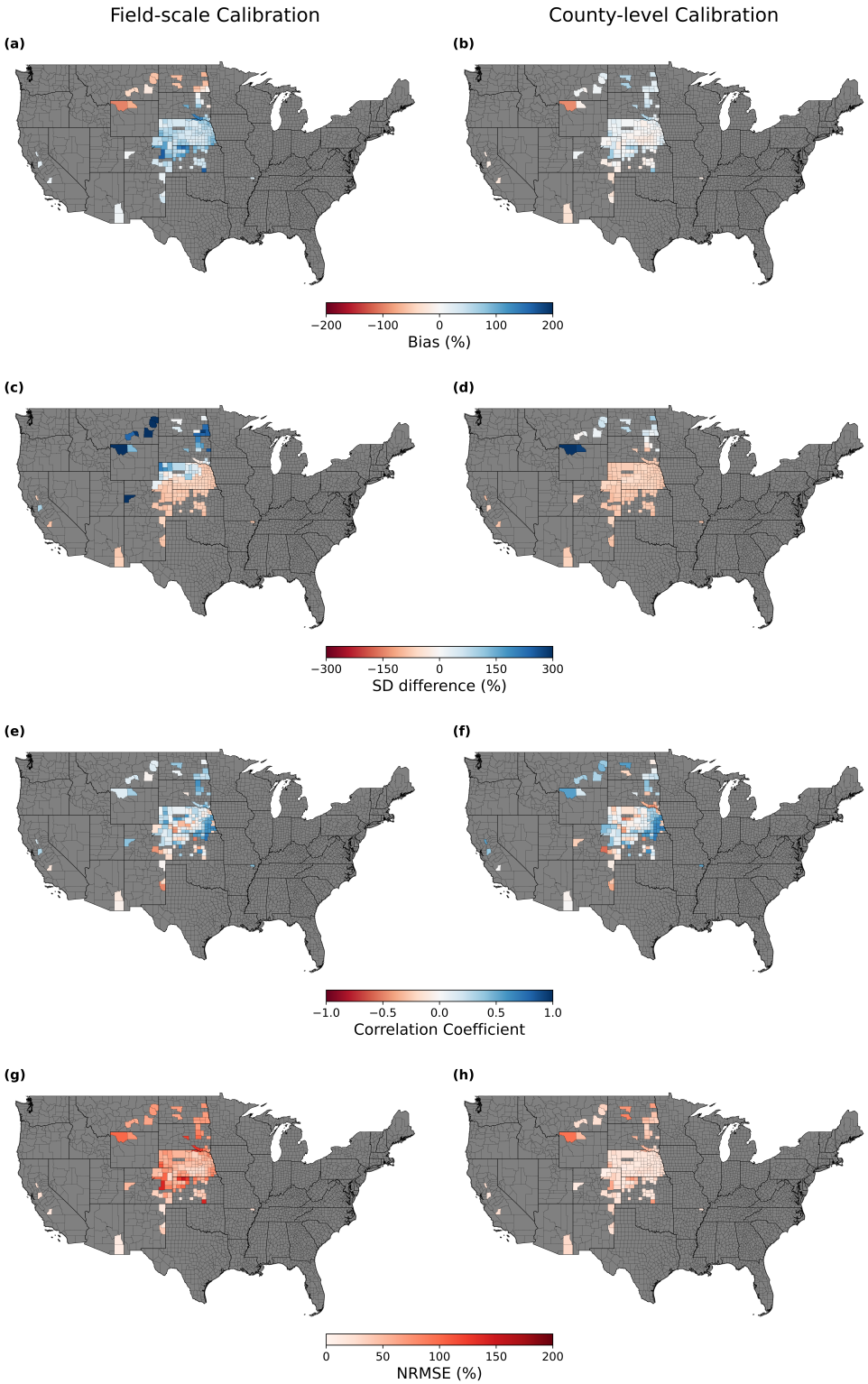

Figure S2: As Figure 2.3 in the main text, but for irrigated counties.

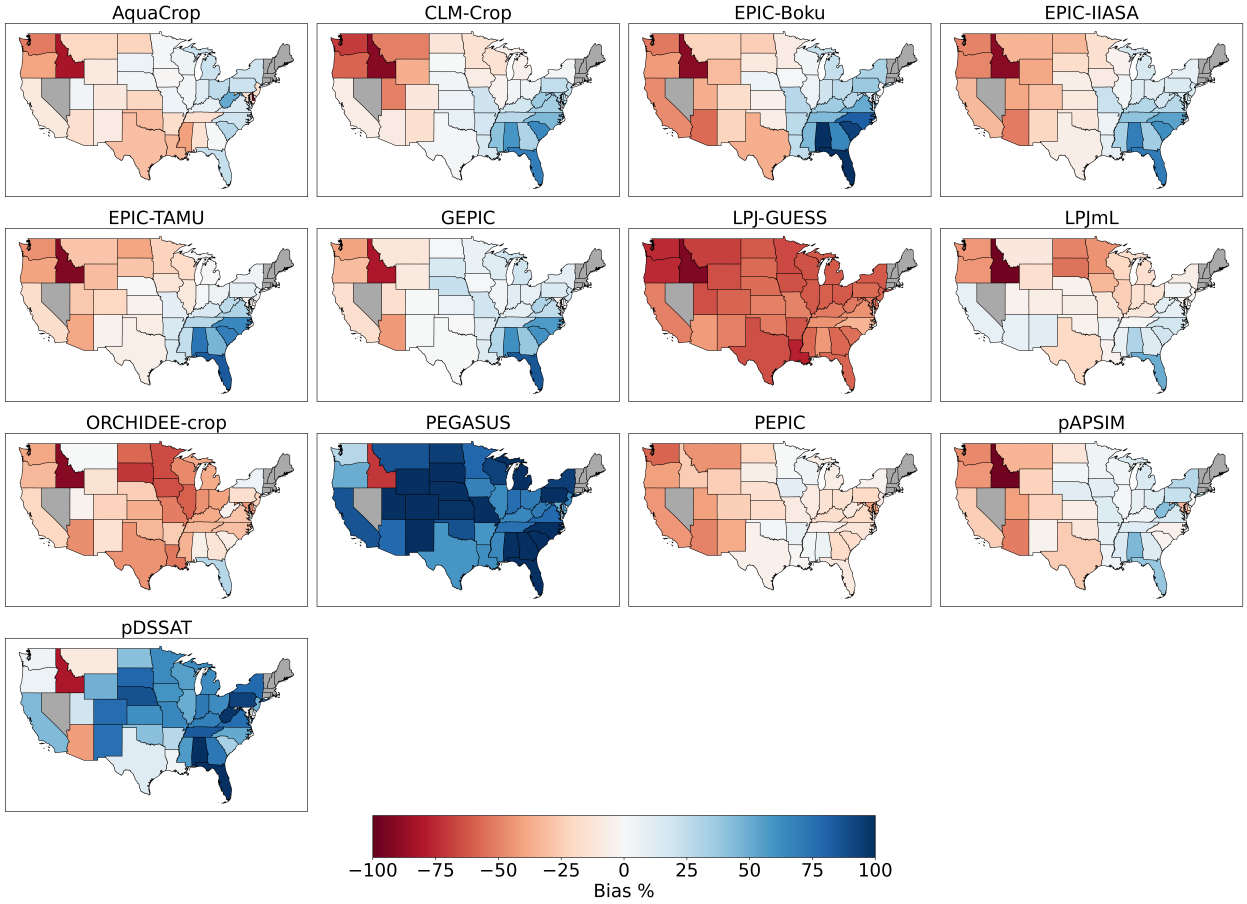

Figure S3: As Figure 2.4 in the main text, but for harmnon harmonization scenario.

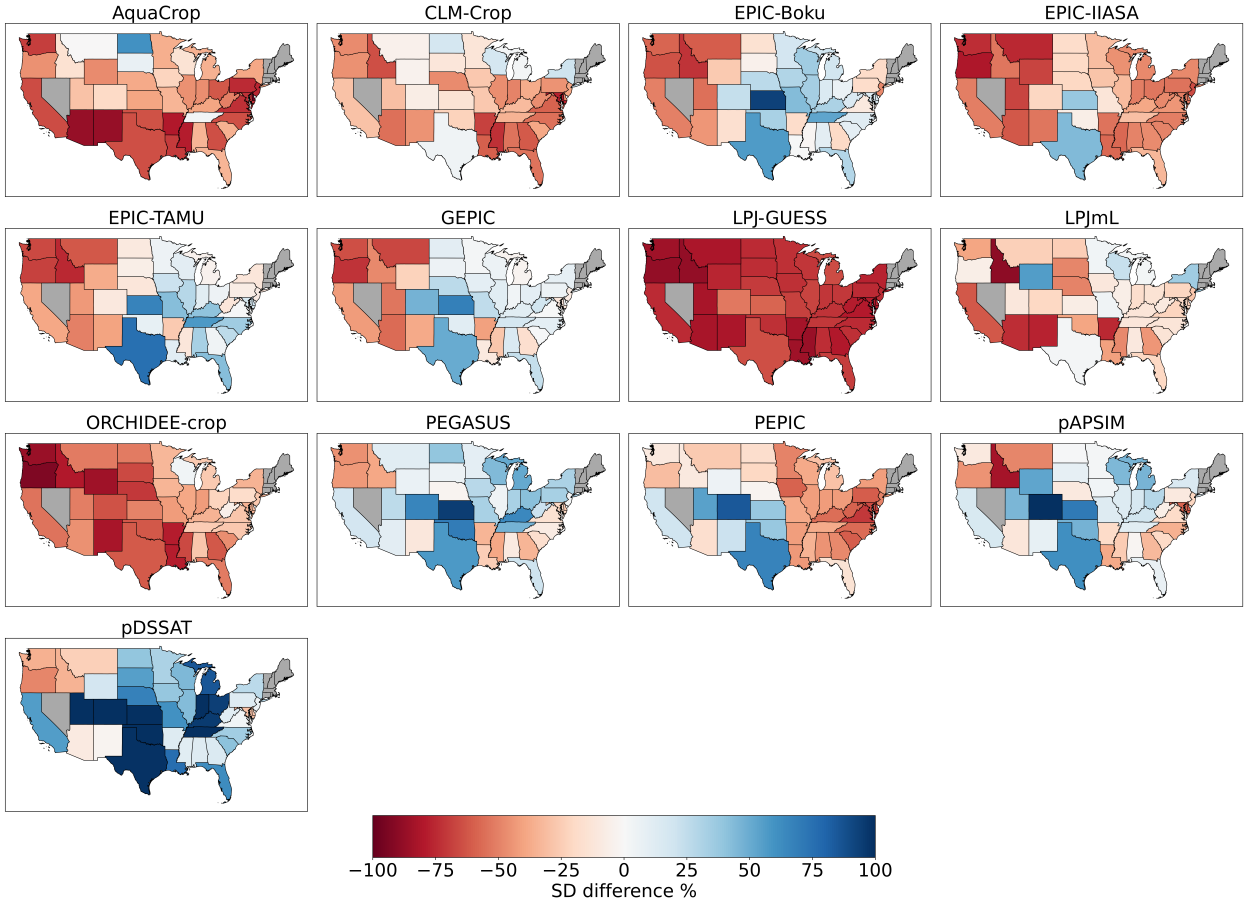

Figure S4: As Figure 2.5 in the main text, but for harmnon harmonization scenario.

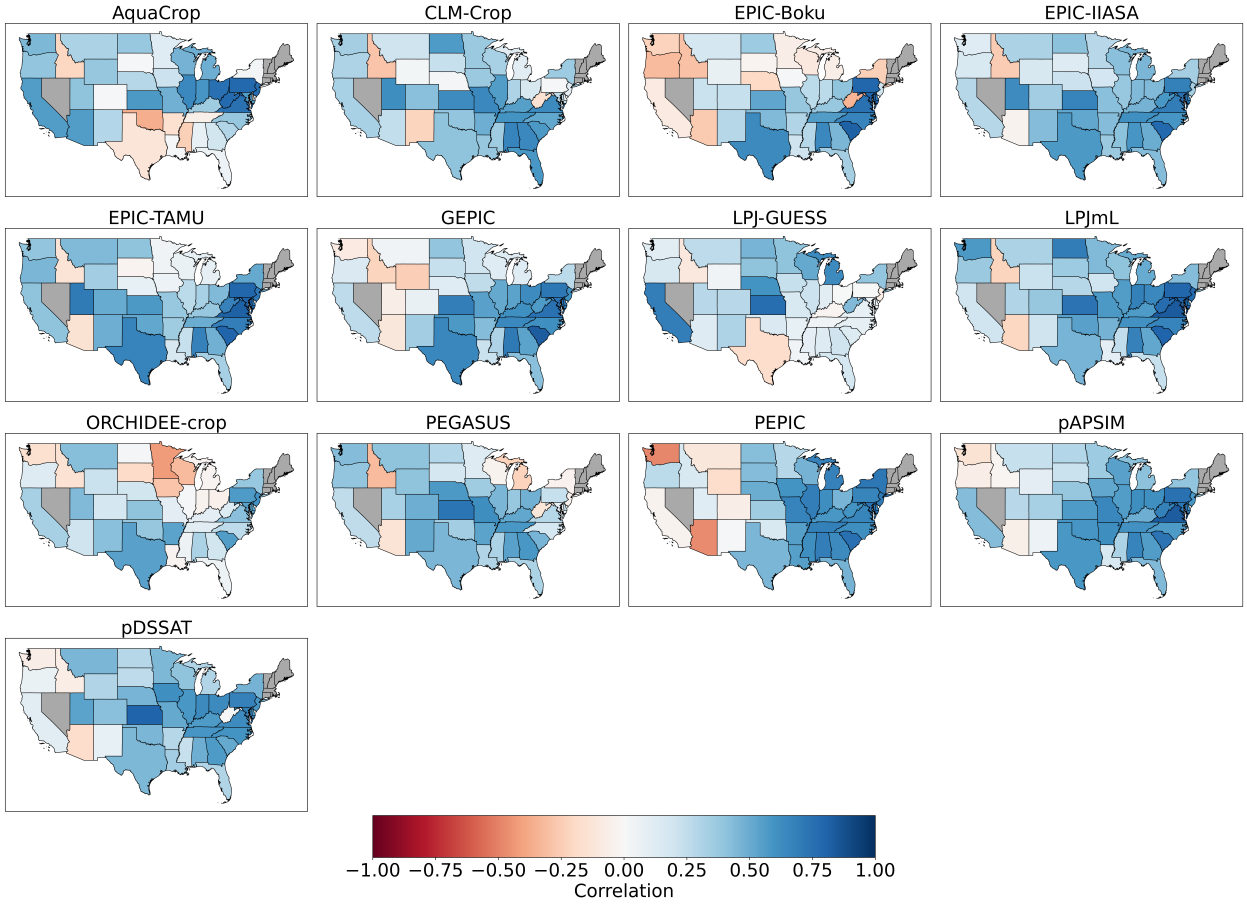

Figure S5: As Figure 2.6 in the main text, but for harmnon harmonization scenario.

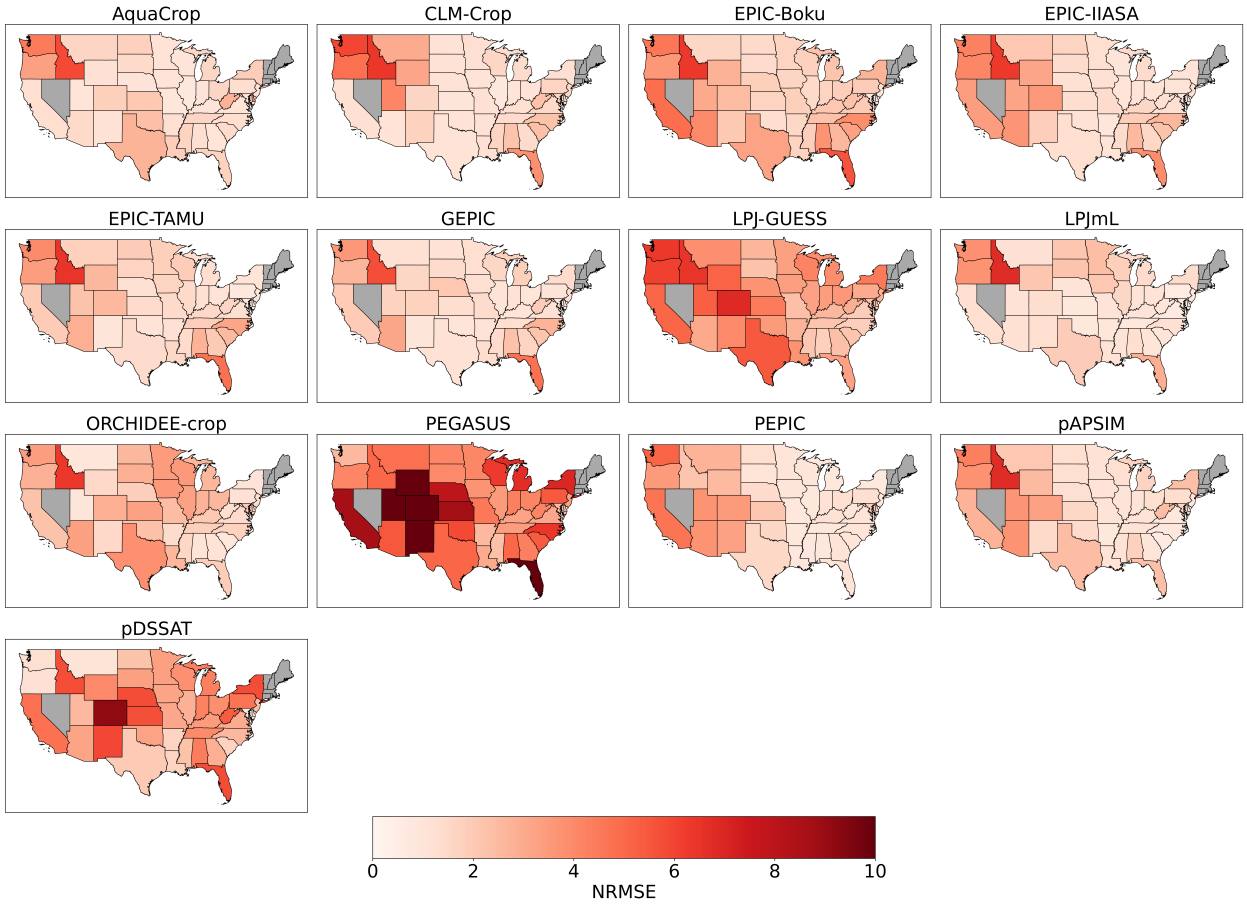

Figure S6: As Figure 2.7 in the main text, but for harmnon harmonization scenario.

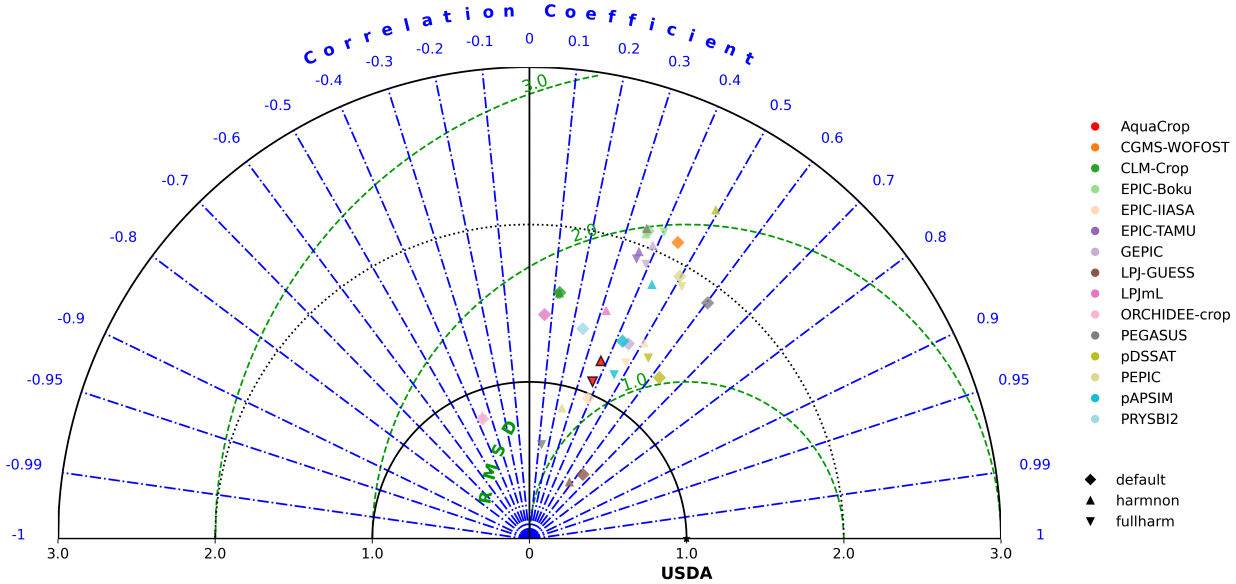

Figure S7: As Figure 2.10 in the main text, but for temporal-dynamic only.

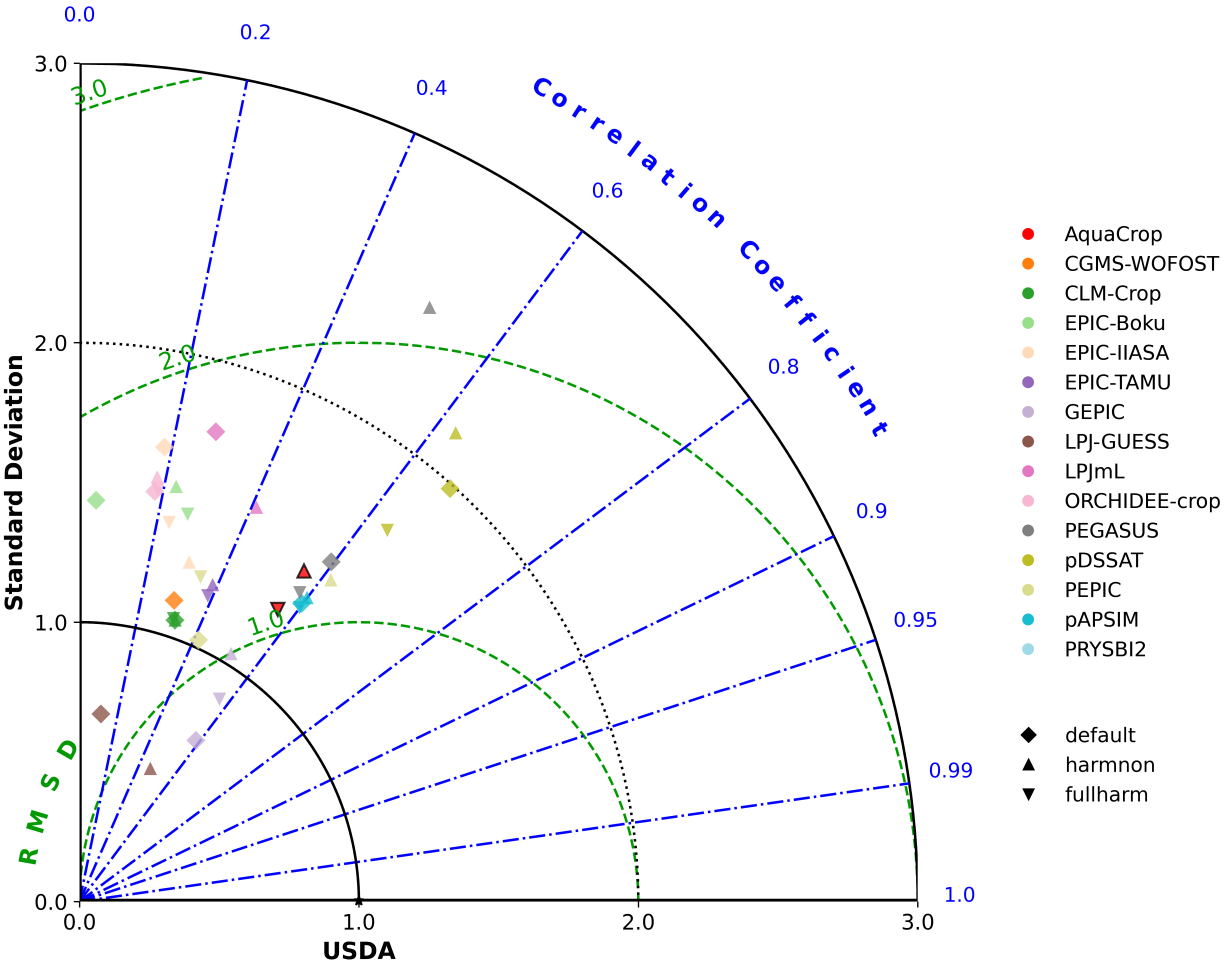

Figure S8: As Figure 2.10 in the main text, but for spatial-dynamic only.

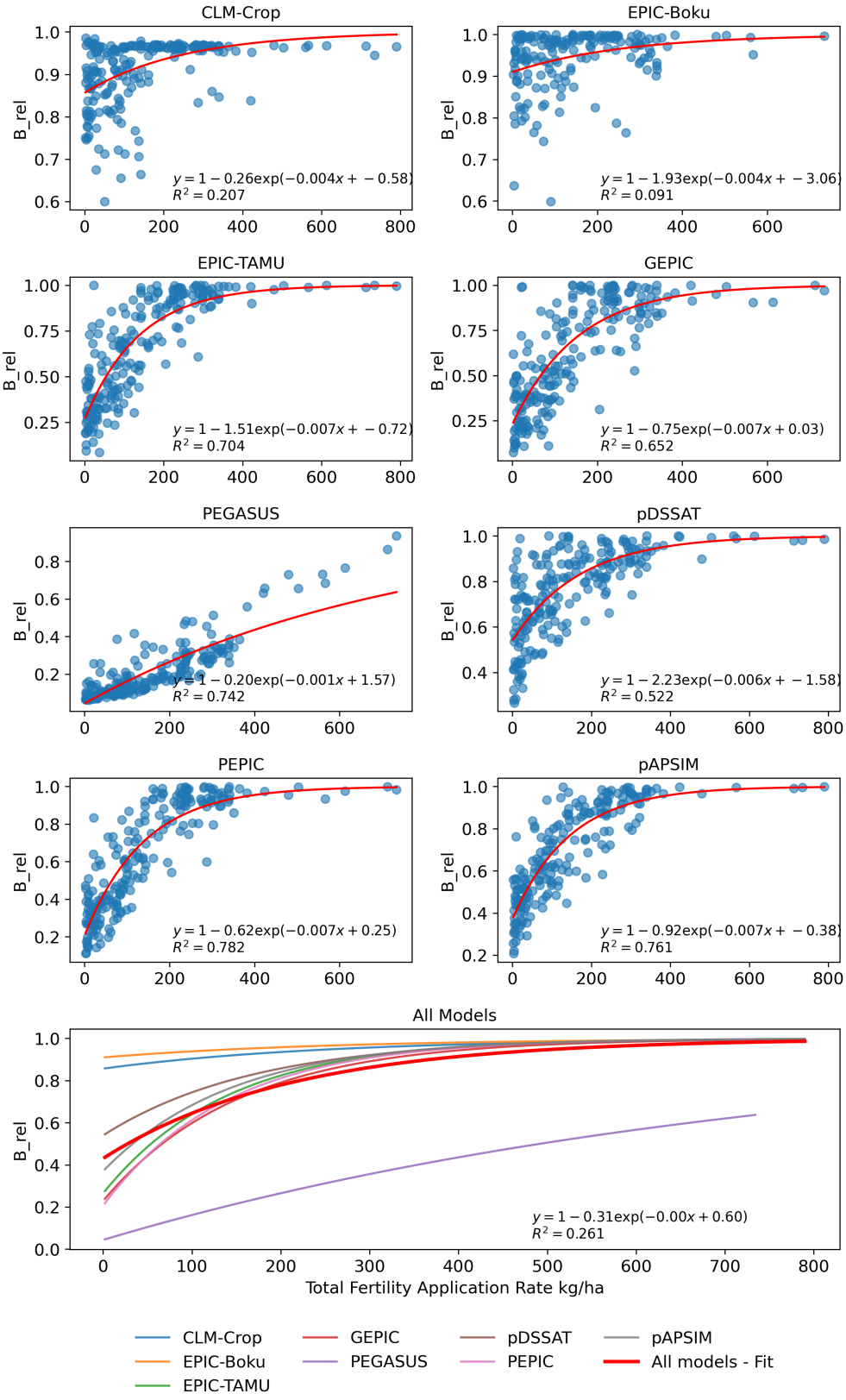

Figure S9: Relationship btween B_rel (ratio of fertility-limited biomass and stress-free biomass under irrigation) and total fertility application rate with Mitscherlich equation fit.

Chapter 3

County-level Projections of U.S. Maize Production under Climate Change: Impacts, Yield Gaps, and Adaptation Strategies

Wu, S., Monier, E. (2025). County-level projections of U.S. maize production under climate change: impacts, yield Gaps, and adaptation strategies. In preparation for Earth's Future.

Abstract

Maize is the most extensively cultivated cereal crop and plays a pivotal role in global food supply. However, its production is highly vulnerable to climate change. In this study, we use a flexible Python-based implementation of AquaCrop to simulate county-level maize yields across the contiguous United States which is the world's leading maize producer. Simulations are driven by a large ensemble of climate projections from 13 CMIP6 global climate models under two scenarios, SSP126 (low warming) and SSP585 (high warming). We examine (1) national-scale and county-level maize productivity changes, (2) yield gaps between rainfed and irrigated systems, (3) the effectiveness of irrigation and fertilization as adaptation strategies, and (4) the potential for relocating production to new regions. We project that U.S. maize production could decline by 17% by the end of the century under SSP585, and by 1% under SSP126. Maize yields across the Corn Belt region are projected to decline by 10–20%, highlighting this region's high vulnerability to climate change. Perfect irrigation and fertilization for existing maize systems can offset yield losses during the mid-century and their combined implementation can increase production by roughly 30%, but under SSP585 neither measure alone fully offsets climate damages. Relocating maize production to northeastern and western counties provides additional potential, although its feasibility is constrained by land protection and water scarcity. Our analysis highlights the importance of water management, nutrient availability, and spatial adaptation for sustaining maize productivity. By providing county-level projections, we generate insights directly relevant to agricultural policy, crop insurance, and farm-level decision-making.

3.1 Introduction

Securing adequate food supplies for the world's population represents one of the most pressing global challenges, particularly in a changing climate (Kang et al., 2009; Wheeler & Von Braun, 2013). Rising temperature, shifting precipitation patterns, more frequent extreme events, and elevated CO_2 concentration are expected to affect both staple and perennial crop production (Hong et al., 2020; Leisner, 2020; Rezaei et al., 2023; Wu et al., 2025). Climate change therefore threatens to exacerbate existing pressures on global food systems, with potentially severe consequences for food security, livelihoods, and agricultural sustainability (IPCC, 2019; Myers et al., 2017; Wheeler &

Von Braun, 2013). As the most extensively cultivated and consumed cereal in the world, maize is a key crop to examine, and understanding climate change impacts on maize yields provides important insights into the future of food production. Globally, maize accounts for nearly 20% of total calories consumed by humans and livestock and is produced in volumes exceeding 1.1 billion metric tons annually (FAO, 2024), underscoring its central role in both food and feed systems worldwide.

Maize has already experienced the effects of rising temperatures and altered precipitation patterns, which have intensified heat and drought stresses on production (Djalovic et al., 2024; K.-H. Kim & Lee, 2023; R. W. Portmann et al., 2009; Vadez et al., 2024). Extreme drought events during the development stage have led to yield reductions of up to 43% in the southeastern U.S., and increasing drought sensitivity has been observed nationwide (Lobell et al., 2020, 2014; Nguyen et al., 2023). Using an ensemble of process-based crop models under two Shared Socioeconomic Pathway (SSP) scenarios (Riahi et al., 2017), Jägermeyr et al. (2021) report that global maize yields will decline by 24.1% under a high-warming scenario (SSP585) and by 6.4% under a low-warming scenario (SSP126) by the end of the century. In contrast, global wheat yields increase by 17.5% and 8.8% under SSP585 and SSP126, respectively. Similarly, C. Li et al. (2025), using statistical crop models, estimate that maize will experience the steepest decline (22.2%) among maize, rice, soybean, and wheat under SSP585. Both studies also highlight that climate change poses greater risks to maize productivity in the United States than elsewhere, with projected yield losses of 26–40% by century's end. As the world's leading maize producer, the U.S. supplies nearly 30% of global maize and generates over 70 billion dollars annually from the crop (FAO, 2024; USDA NASS, 2025). Given maize's central role in U.S. food, feed, and biofuel systems (García-Lara & Serna-Saldivar, 2019), these findings underscore the urgent need to assess climate impacts on maize productivity and to evaluate practical adaptation strategies.

Water availability is central to maize responses under climate change. Water shortage constrains physiological processes such as germination, leaf growth, transpiration, and grain filling (K.-H. Kim & Lee, 2023). Yang & Wang (2023) show that water stress is the primary driver of projected maize yield losses in the Corn Belt before mid-century, while Caparas et al. (2021) estimate that the probability of water stress–induced yield failure will more than double by mid-century. Irrigation can offset these risks. For example, in the Ogallala Aquifer region, adding 100 mm of irrigation reduces maize heat sensitivity by 7.6% (L. Zhang et al., 2025). Irrigation also helps stabilize yield

variability (Irmak et al., 2022), although Partridge et al. (2023) argue that irrigation costs may exceed benefits across most U.S. maize croplands by mid-century. Assessing the yield gap between rainfed and irrigated systems under future climates therefore provides important insight into both the potential and the limitations of irrigation as an adaptation measure.

Most previous studies projecting U.S. maize yields under climate change have mainly generated gridded simulations (Jägermeyr et al., 2021; Partridge et al., 2023), while county-level projections remain limited to a few states (Johnston et al., 2015; Urban et al., 2015). Yet most relevant datasets—including yields, harvested areas, crop risk management data, water use, and irrigation amounts—are primarily reported at the county scale (Dieter et al., 2018; Elias et al., 2018; National Academies of Sciences et al., 2023; Ruess et al., 2023). County-level projections therefore provide critical advantages for integrating with these datasets in both research and decision-making. Moreover, existing work largely focuses on current maize-growing areas, overlooking regions that may become suitable for cultivation under future climates (Beyer et al., 2022).

To address these gaps, we use the gridded, Python-based implementation of AquaCrop (Wu & Monier, 2025), originally developed by the Food and Agriculture Organization for field-scale simulations (Steduto et al., 2009). AquaCrop explicitly represents crop–atmosphere–soil interactions and simulates canopy cover, transpiration, biomass, and yield at a daily time step while accounting for water, temperature, fertility, and salinity stresses as well as management practices such as irrigation. Because the default maize parameterization in AquaCrop is based on conditions at the University of California, Davis, which do not adequately represent the diverse environments of U.S. maize-growing regions (Hsiao et al., 2009), we apply parameters calibrated for counties with substantial maize production but distinct climates and geographical locations. Using this framework, we provide county-level projections of U.S. maize yields under future climates, quantify yield gaps between rainfed and irrigated management, and assess the potential and limitations of irrigation as an adaptation strategy.

3.2 Materials and Methods

3.2.1 AquaCrop

AquaCrop simulates crop growth on a daily time step by explicitly modeling crop physiological processes and biophysical interactions among the crop, climate, soil, and water management (Steduto et al., 2009). Daily yield estimation follows four major steps: (1) calculating the canopy cover (CC) fraction based on the accumulation of growing degree days (GDD); (2) calculating crop transpiration (Tr) by multiplying reference evapotranspiration (ET_o) by a crop coefficient (K_{ctr}) that is proportional to canopy cover; (3) normalizing biomass water productivity based on atmospheric CO_2 concentration and multiplying it by the accumulation of Tr/ET_o to obtain above-ground biomass; and (4) multiplying the harvest index (HI) with accumulated biomass to estimate attainable yield. In this study, we use a gridded, Python-based implementation of AquaCrop. A detailed description of the calculation procedure and model setup is provided in (Wu & Monier, 2025).

Crop responses to stresses

AquaCrop accounts for multiple stresses, including water shortage, waterlogging, extreme temperature, salinity, and limited fertility. Water stress inhibits canopy development and accelerates canopy decline. Stress-induced stomatal closure reduces transpiration and constrains root expansion. Harvest index also decreases when water stress impairs reproductive processes such as pollination or reduces stomatal conductance (Raes et al., 2009). AquaCrop incorporates a detailed soil water balance, simulating daily water fluxes at each soil depth and computing stress coefficients for canopy expansion, canopy senescence, pollination, and stomatal closure based on corresponding thresholds. For temperature stresses, harvest index is adjusted for both cold and heat extremes, with cold stress additionally decreasing transpiration. Canopy cover and transpiration are both reduced under salinity stress, while fertility stress primarily limits canopy development and biomass accumulation.

Crop response to elevated CO_2

AquaCrop considers both the reduction in transpiration and the enhancement of biomass production under elevated CO_2 concentrations. Elevated CO_2 decreases crop stomatal conductance, thereby

reducing transpiration. AquaCrop adjusts the crop coefficient for maximum transpiration (Kc_{Trx}) to capture this effect. At the same time, water productivity is increased to account for the fertilization effect of higher CO_2, which enhances biomass accumulation.

Calibration

The parameters used in this study are calibrated with county-level maize yield data from the United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) (USDA NASS, 2023). The calibration includes 8 irrigated counties and 9 rainfed counties strategically selected to represent diverse environmental conditions of maize-growing areas across the contiguous U.S. under rainfed and irrigated management. Wu & Monier (2025) demonstrates that the county-level calibration significantly improves AquaCrop's performance for large-scale simulations.

3.2.2 Model input

To investigate the impacts of climate change coupled with water management on maize yields across the contiguous U.S., we use a large ensemble of high-resolution global climate projections: the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) based on the Coupled Model Intercomparison Project Phase 6 (CMIP6) (O'Neill et al., 2016; Thrasher et al., 2022). The NEX-GDDP-CMIP6 dataset is derived through statistical downscaling of 35 General Circulation Model (GCM) outputs from CMIP6 for one historical experiment and four Shared Socioeconomic Pathways (SSP) experiments (Riahi et al., 2017). The downscaling process employs the daily Bias Correction Spatial Disaggregation (BCSD) method (Wood et al., 2004) using the Global Meteorological Forcing Dataset for Land Surface Modeling (GMFD) as observational input (Sheffield et al., 2006). NEX-GDDP-CMIP6 provides daily climate variables at a spatial resolution of 0.25 under both historical (1950–2014) and future (2015–2100) scenarios, including SSP126, SSP245, SSP370, and SSP585. In this study, we simulate maize yields with climate projections from 13 GCMs under SSP126 and SSP585, which represent low- and high-warming scenarios, respectively.

AquaCrop determines soil hydraulic properties based on soil texture (proportions of sand, silt, and clay). We employ gridded soil data from the Harmonized World Soil Database (Fischer et al., 2008), consistent with previous large-scale crop yield simulations (Jägermeyr et al., 2021; Müller et al., 2019). Irrigated and rainfed harvested areas for maize are obtained from the MIRCA2000

dataset (F. T. Portmann et al., 2010), which provides gridded cropland area data for 26 crops. The growing period for each county is identified using typical planting and harvesting dates from Sacks et al. (2010). Nutrient input is represented by gridded fertilizer application rates produced by the GGCMI Phase 1 project (Müller et al., 2019).

3.3 Results

3.3.1 Future U.S. maize production change

Climate change is projected to negatively affect maize production in the U.S. under both SSP126 and SSP585 scenarios (Fig. 1). Yields under purely rainfed and perfectly irrigated conditions are simulated for each county across the contiguous U.S. We aggregate simulated county-level yields (for counties that reported maize yields to the USDA during 2000–2019) to estimate U.S. national maize production, weighted by irrigated and rainfed harvested areas from MIRCA2000. Compared with the historical period (2000-2019), U.S. maize production declines by roughly 1% (multi-model median across 13 GCMs) with ranges between –9% to 2% under the low-warming scenario (SSP126), and by 17% with a confidence interval of –27% to –9% under the high-warming scenario (SSP585) by the end of the century (2080–2099). Under SSP126, which represents an optimistic scenario assuming atmospheric CO_2 concentrations gradually decline after mid-century (Tebaldi et al., 2021), simulated decreases in maize production are similar for the mid-century period (2040–2059) and the end-of-century period (2080–2099). In contrast, production under SSP585 demonstrates notably larger reductions than SSP126 after the mid-2070s. All models agree on the production decline by 2080-2099, indicating robust directionality under high warming. In addition, 8 of 13 GCM-driven simulations include at least one year in the latter half of the century with national production losses exceeding 40%. The decline in rainfed maize production (17.8%) is larger than that in irrigated maize production (13.7%), and rainfed simulations demonstrate more intense uncertainty which indicates the large uncertainty in precipitation projections across GCMs

3.3.2 County-level yield change

During the middle of the century, most counties in the Corn Belt region experience yield losses between 0 and 10%, while the southeast coast achieves modest yield gains of up to 10% (Fig. 2).

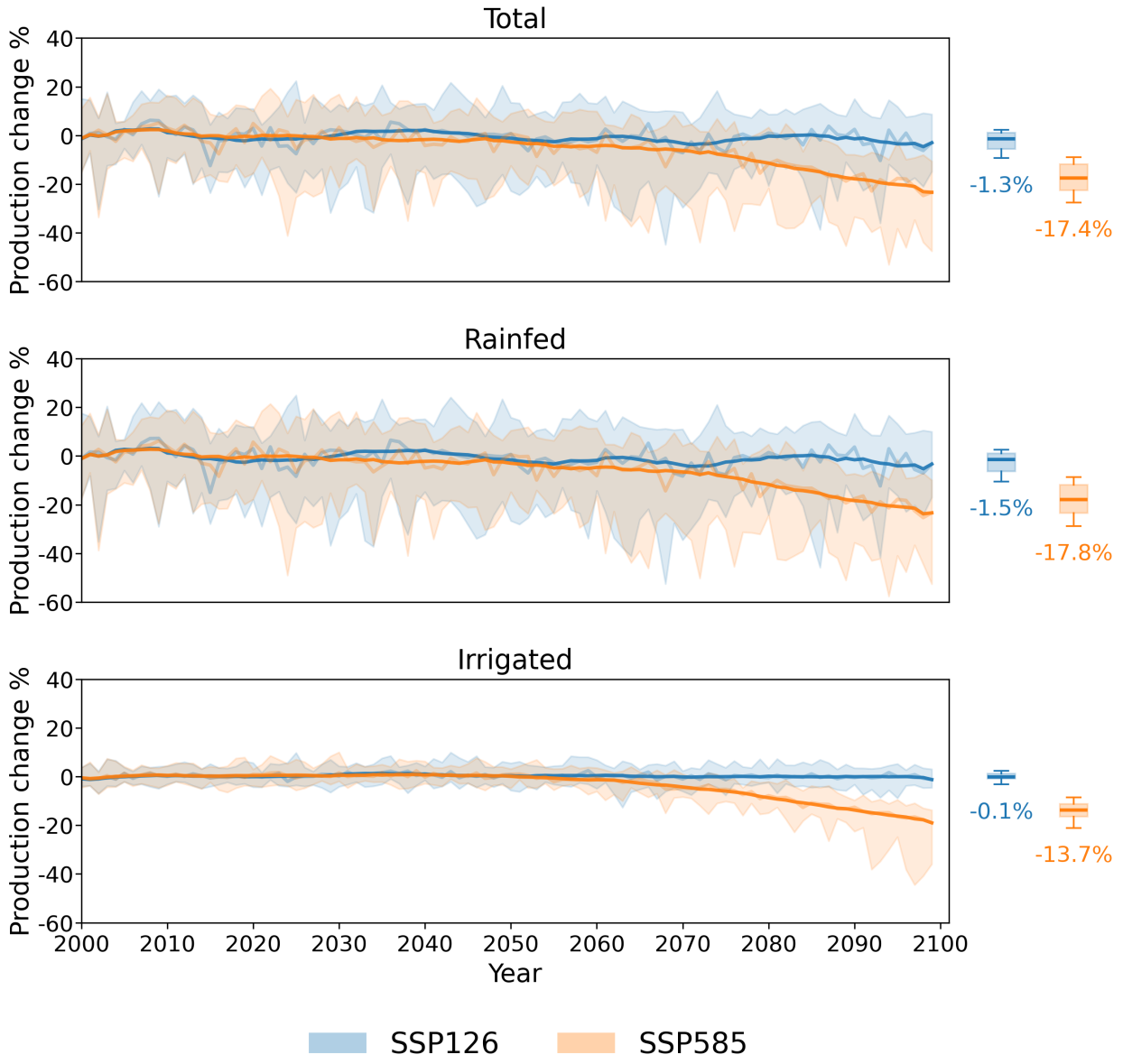

Figure 3.1: Time-series of yearly U.S. national maize production median changes (%) (and 10yr moving average) relative to the mean nationwide production during the historical period (2000-2019) under SSP126 (**blue**) and SSP585 (**orange**) climate scenarios. Shadings represent the ranges of simulations with 13 GCMs. Box plots show the distributions of end-of-century (2080-2099) production changes, followed by multi-model median of productions changes.

Under both climate scenarios, yield gains are projected for the Rock Mountains region. By the end of the century, county-level yield changes under SSP126 show similar patterns and magnitudes to mid-century results (Fig. 3). In contrast, SSP585 leads to nationwide yield declines, particularly in northern states and parts of Texas. In the Corn Belt region, end-of-century outcomes shift toward larger losses, with yield declines spanning –10 to –30%. Counties near the Rocky Mountains continue to benefit, which might result from the inclusion of climate data over the non-agricultural

mountainous areas.

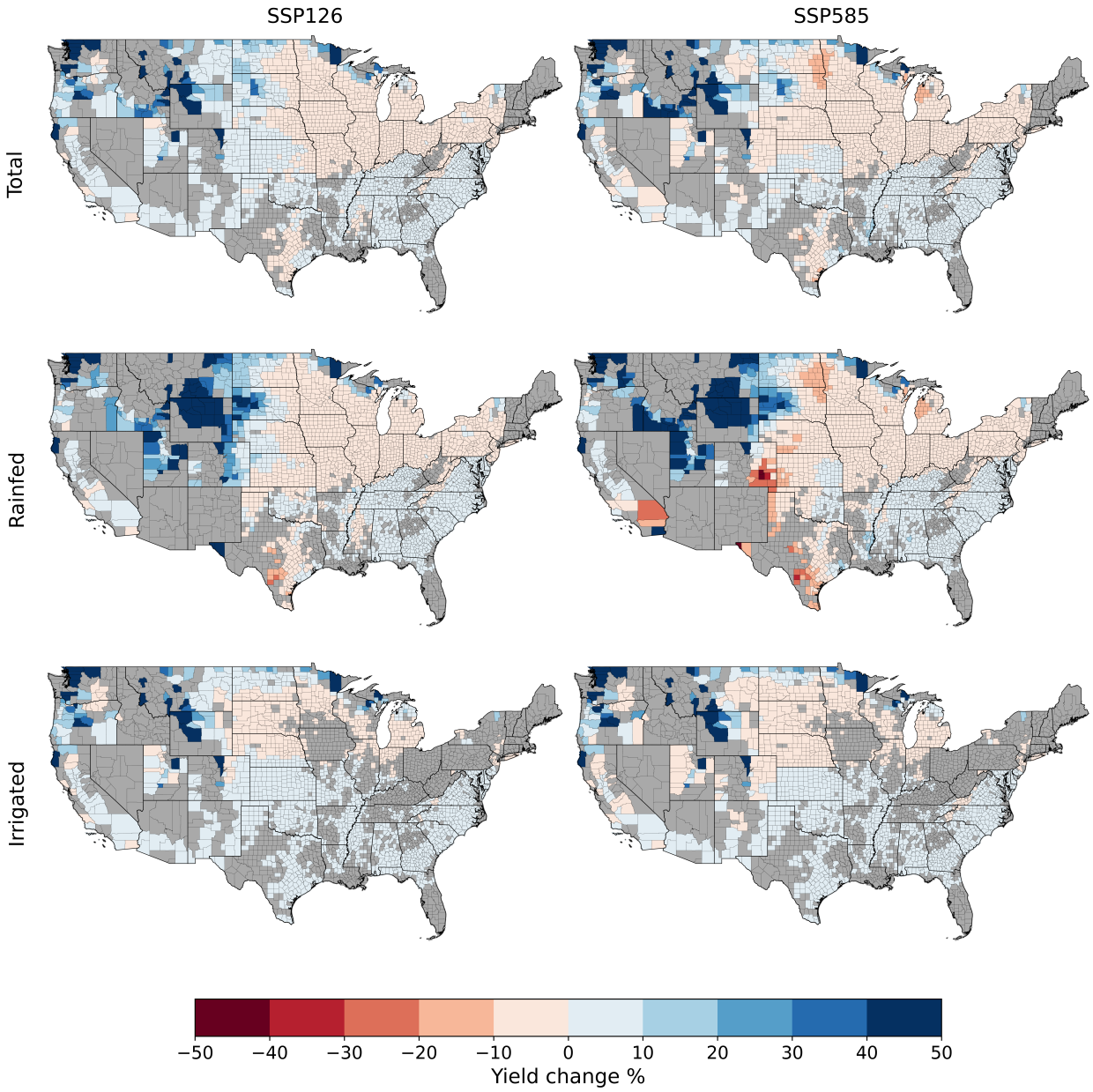

Figure 3.2: U.S. maps showing county-level multi-model mean yield change (%) for mid-century period (2040-2059). Combined yield change is weighted by irrigated and rainfed harvested areas. Rainfed (irrigated) yield changes are plotted for counties with rainfed (irrigated) cropland area exceeding 1% of total area.

3.3.3 Effects of adaptation strategies

To explore potential pathways for mitigating adverse climate impacts on the maize industry, we examine the effects of irrigation and fertilization (Fig. 4). Perfect irrigation of historically rainfed croplands not only offsets climate change impacts by mid-century but also increases national maize

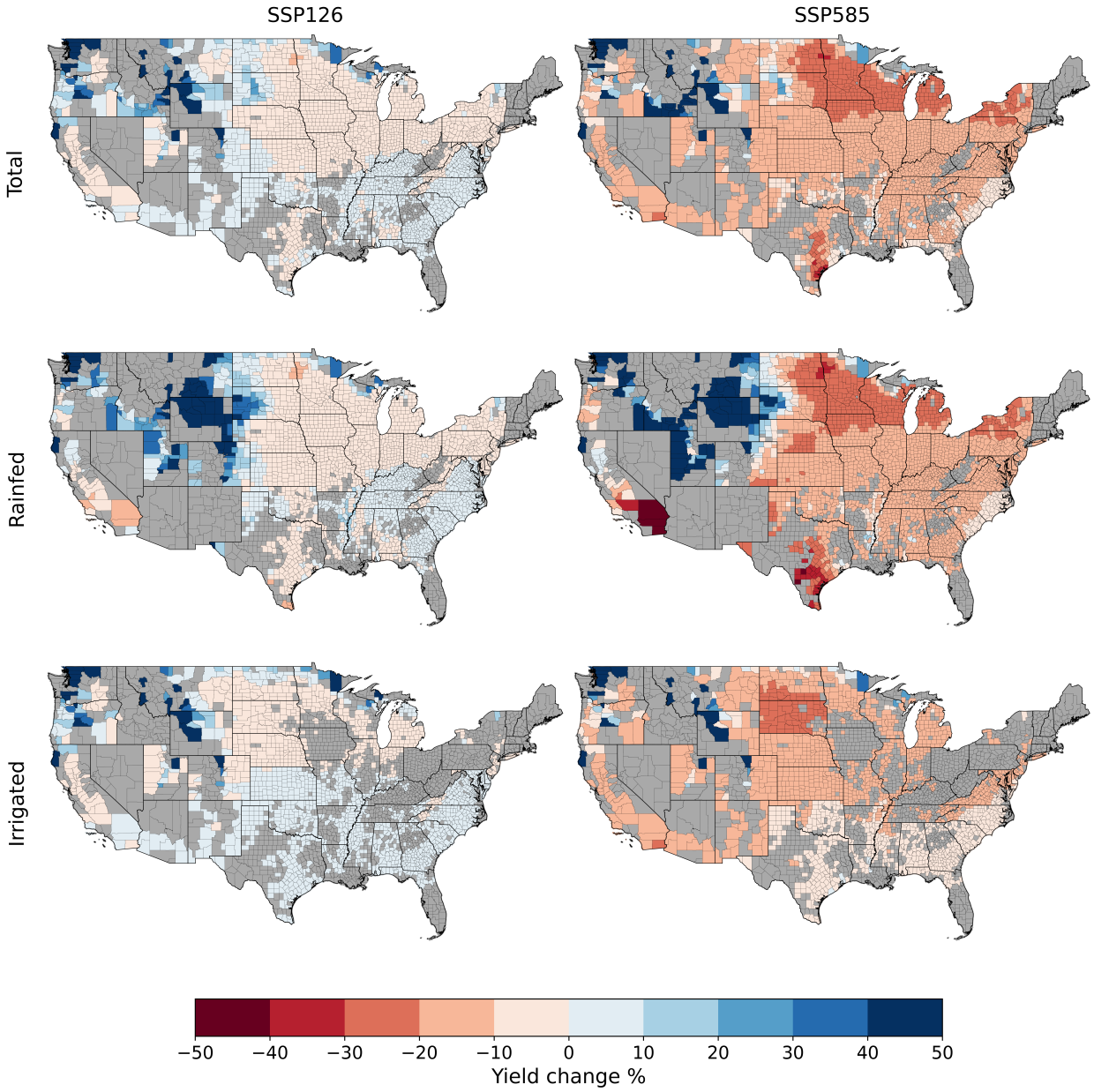

Figure 3.3: Same as Fig. 2 but for end-of-century period (2090-2099).

production by about 14% and 13% under SSP126 and SSP585, respectively. Assuming perfect fertilization in all maize-growing counties results in increases of 10% (SSP126) and 8% (SSP585). Notably, the combined effect of irrigation and fertilization can increase maize productivity by roughly 28%. By the end of the century, both strategies still provide yield benefits under SSP126. However, under SSP585 neither irrigation nor fertilization alone fully offsets production losses, but their combined effect does offset the climate damages and even yields a 9% increase.

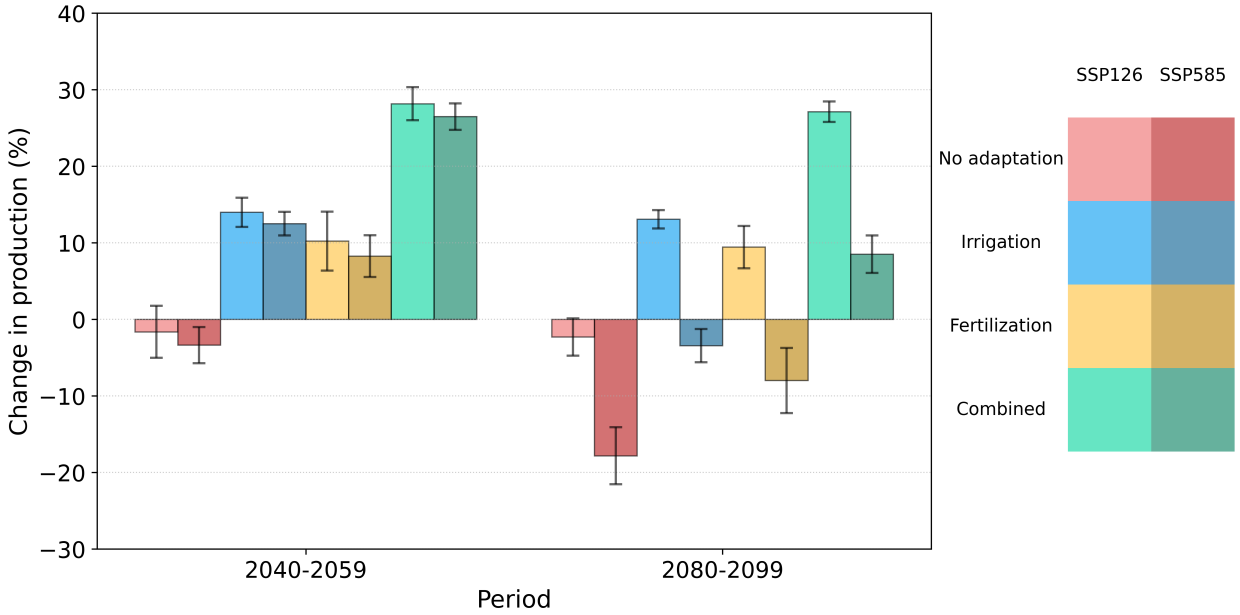

Figure 3.4: Changes (%) of U.S. national maize productions compared with the mean historical production for 4 adaptation scenarios: no adaptation (**red**), perfectly irrigating all rainfed maize croplands (**blue**), perfectly fertilizing all maize-growing counties (**yellow**), and combining both irrigation and fertilization (**green**). Light (dark) bars represent simulations under SSP126 (SSP585). Error bars represent the 95% confidence interval.

3.3.4 Counties with high yield potential

In addition to applying adaptation measures to existing maize-growing areas, cultivating maize in new counties with high yield potential offers another strategy for building resilience in the U.S. maize industry (Fig. 5, Fig. 6). Under purely rainfed conditions, counties in the northeastern U.S. provide higher simulated maize yields than 95% of current maize-growing counties by mid-century under both scenarios. Some southeastern counties also show relatively high yield potential. By the end of the century, however, only a limited number of these counties exceed 50% of current production regions under SSP585. With unlimited water supply, western counties demonstrate substantial yield potential by mid-century, but these regions typically face limited water availability, constraining irrigation feasibility. Under SSP585, states such as Arizona cannot outperform most current maize-growing regions during 2090–2099, even under perfect irrigation.

3.4 Discussion and Conclusion

As the most extensively cultivated cereal crop in the world, maize is a crucial component of the global food supply, supporting both direct human consumption and livestock production. Assessing

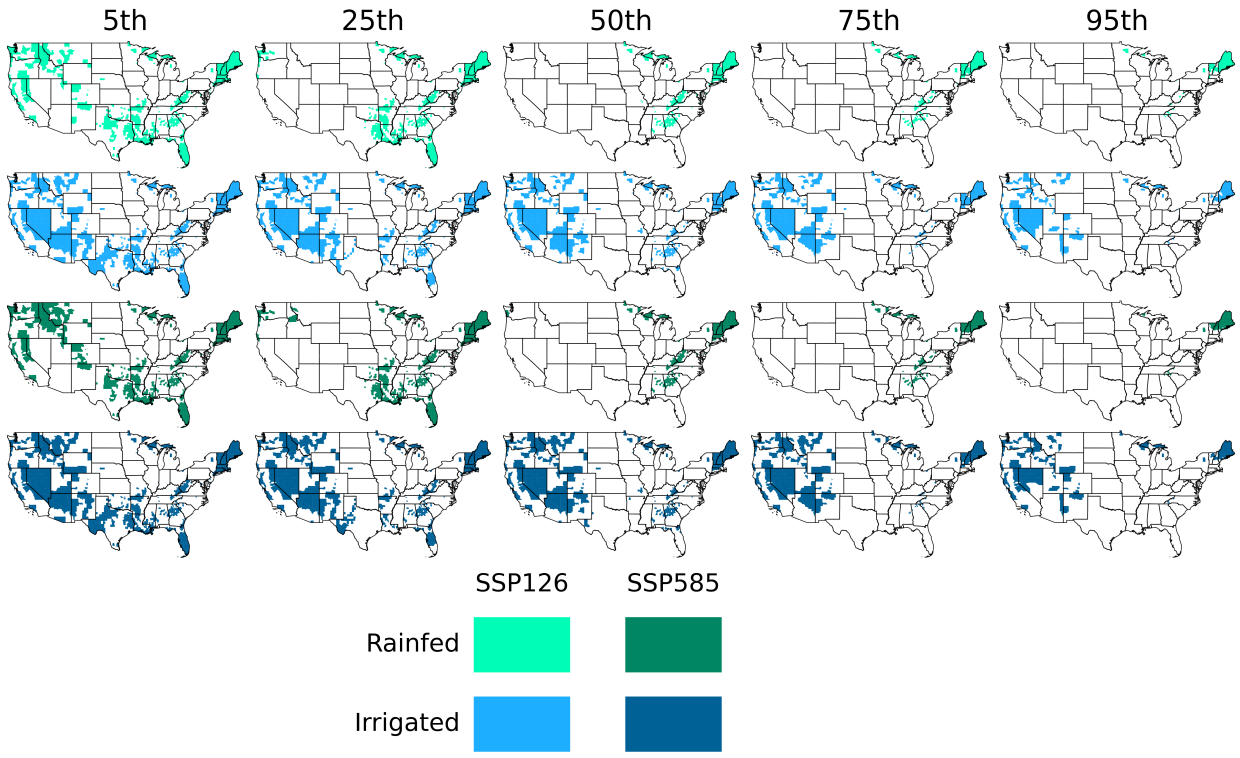

Figure 3.5: U.S. maps showing counties that did not report maize yield to the USDA during the historical period but are simulated to achieve yields exceeding the 5th, 25th, 50th, 70th and 95th percentile levels of mid-century (2040-2059) yields of historically maize-growing counties. **Green** (**blue**) represents simulated rainfed (irrigated) yield under SSP126 (light) and SSP585 (dark) scenarios.

the impacts of climate change on maize production and exploring adaptation pathways are essential for ensuring future food security. In this study, we employ a gridded, Python-based implementation of the process-based crop model AquaCrop to simulate county-level maize yields in the U.S., the world's largest maize producer. Using a large ensemble of high-resolution climate projections from 13 CMIP6 GCMs under two representative scenarios, we examine: (1) how climate change affects U.S. maize productivity, (2) the extent of yield gaps between rainfed and irrigated systems, (3) the potential benefits of adaptation strategies such as irrigation and fertilization, and (4) the opportunities and constraints of relocating maize production to new regions.

Our simulations project a multi-model median decline of 18% in U.S. national maize production by the end of the century under the high-warming scenario (SSP585), and a smaller decline of 1% under the low-warming scenario (SSP126). These results are broadly consistent with prior global-scale analyses, including a multi-model ensemble of 12 global gridded crop models (GGCMs) driven by CMIP6 projections (Jägermeyr et al., 2021) and statistical modeling studies incorporating

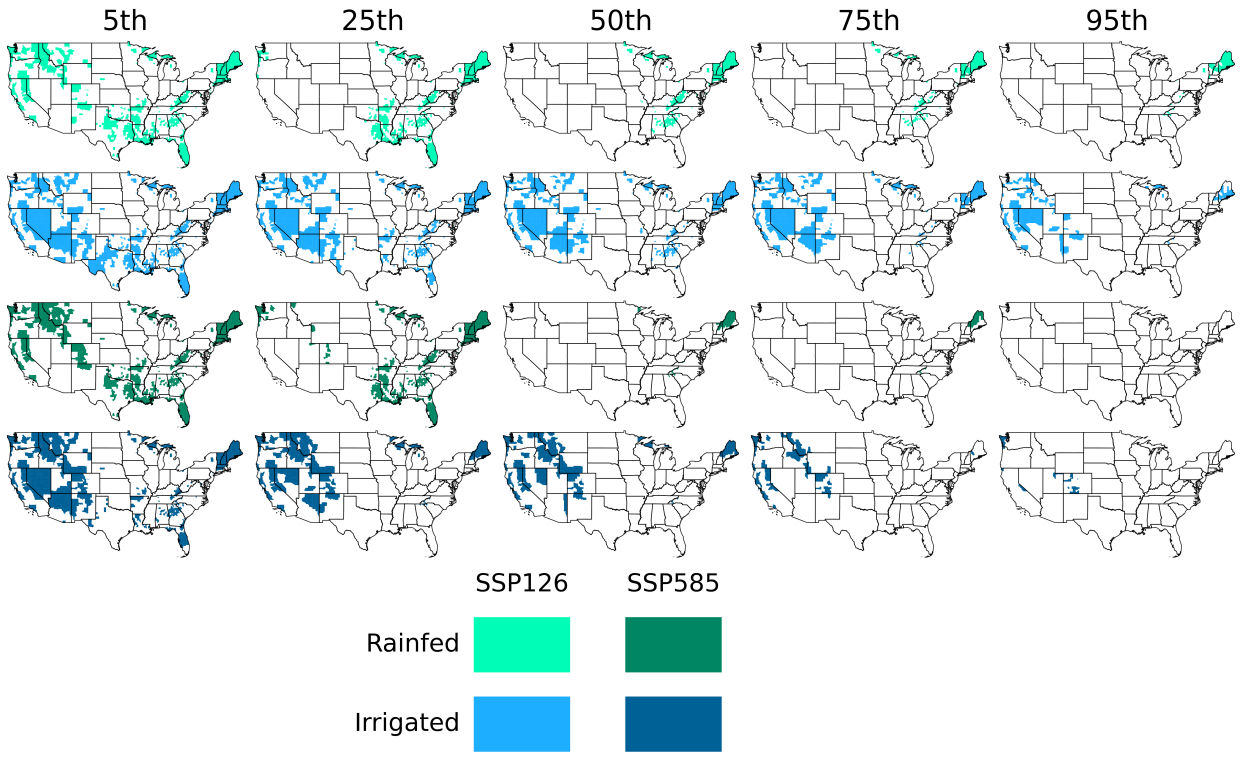

Figure 3.6: Same as Fig. 4 but for end-of-century period (2080-2099) period.

23 GCMs (C. Li et al., 2025). We also find substantial yield reductions of 10–20% in the Corn Belt, confirming the region's high vulnerability to warming despite differences in magnitude compared to other studies. For example, Hultgren et al. (2025) report losses of up to 40%, while our estimates are smaller but still confirm significant regional risk.

To support the development of adaptation strategies, we quantify production changes in response to irrigating all maize-growing areas and applying optimal fertilization. Our results show that either measure alone offsets climate change damages and even generates yield gains during mid-century. In combination, these strategies increase national maize production by nearly 30%. However, under the warmest conditions (end-of-century SSP585), only their joint implementation fully counteracts negative impacts. Similarly, Ren et al. (2025) find that the integrated benefits of irrigation and fertilization outweigh associated costs for more than 80% of U.S. maize production, although long-term soil fertility loss may increase fertilizer requirements and expenses (Jang et al., 2021). Moreover, irrigation expansion must avoid depleting freshwater resources or harming aquatic ecosystems (Blann et al., 2009). Enhancing water-use efficiency through improved irrigation scheduling and infrastructure modernization is therefore essential to ensure sustainable intensification. We also

evaluate the potential for relocating maize production to regions with higher future yield potential. Our results suggest that northeastern counties achieve higher productivity than most current maize-growing counties by mid-century under both SSP126 and SSP585, while some western counties show significant potential under sufficient irrigation. Yet relocation faces important constraints: for example, some high-yield areas in Colorado overlap with national forests, precluding large-scale cultivation, while water scarcity limits the feasibility of expanding irrigation in western states such as Arizona (Warziniack & Brown, 2019).

While our work advances understanding of U.S. maize production under climate change, several limitations remain. First, our analysis relies on a single crop model, and thus does not capture structural and parametric uncertainties inherent in multi-model ensembles. Previous studies show that crop model differences can introduce uncertainties as large as those from climate projections themselves (Jägermeyr et al., 2021; Rosenzweig et al., 2014). Second, our harvested area and crop calendar inputs rely on MIRCA2000 and the dataset of Sacks et al. (2010), which represent conditions around the year 2000. These may not fully reflect more recent changes in cropping patterns, irrigation shares, and growing periods. Incorporating updated datasets such as MIRCA-OS (Kebede et al., 2025), which provides harvested areas and crop calendars for 2000–2015, improves accuracy. Third, our adaptation experiments assume perfect irrigation and fertilization across all counties, which is unlikely to be achievable in practice due to economic, environmental, and social constraints. Future studies should integrate biophysical feasibility with economic and policy considerations, particularly regarding water allocation trade-offs across agricultural, municipal, and ecological sectors. Additionally, uncertainties in the CO_2 fertilization effect, as well as interactions with pests, diseases, and technological change, remain unrepresented but could strongly influence future maize outcomes. Finally, we acknowledge that our method to average climate variables for each county can be improved by incorporating the Cropland Data Layer dataset (Boryan et al., 2011) to exclude climate information over areas that are unlikely to cultivate maize such as foothills and mountains. The inclusion of climate data from high-elevation areas can reduce the county-level average temperature especially in those large-size counties near the Rock Mountain regions, which explains the yield gains in this area under cliamte change.

In conclusion, our study shows that climate change poses a substantial threat to U.S. maize production, with national declines of up to 20% projected by the end of the century under high

warming. By explicitly revisiting the objectives, we demonstrate that: (1) climate change reduces maize productivity nationally and in the Corn Belt; (2) irrigation narrows the yield gap between rainfed and irrigated systems; (3) irrigation and fertilization, particularly in combination, provide substantial though incomplete mitigation; and (4) relocating maize to new high-potential regions offers opportunities but faces serious constraints. Together, these findings underscore that sustaining U.S. maize productivity in a changing climate requires integrated strategies that combine biophysical adaptation, resource-efficient management, and supportive policies. Importantly, our county-level results provide critical insights for policy design, crop insurance programs, and farm-level decision-making by aligning directly with the scales at which agricultural management and risk assessments occur. Future work linking crop modeling with water resource planning, socio-economic analysis, and regional adaptation strategies is essential for developing robust pathways that safeguard both food security and environmental sustainability.

Conclusion

This dissertation aims to advance the understanding of how climate change affects both perennial and annual cropping systems and how innovation and adaptation can mitigate those impacts. The first study develops a novel statistical framework that links agro-climate indices (ACIs) to the phenological stages of a major perennial crop (California almonds) to identify impact pathways and inform adaptation. It shows that climate change can substantially reduce California almond yields by the end of the century, with projected median losses of 17% under SSP245 and 49% under SSP585 if current management remains unchanged. Negative impacts are more severe in the southern Central Valley, while innovation gains show heterogeneity in northern counties and are overall lower there. This study provides insights into the direct impact of climate change on each physiological growth stage and thus enables the identification of adaptation strategies to minimize future yield losses that current crop modeling studies do not. It reveals that the dominant damaging ACIs are daily minimum temperature during the bloom stage, especially under mild warming by mid-century or under the lower warming scenario SSP245, as well as humidity levels during the bloom stage (Bloom_Humidity) and the exposure to temperature higher than 35°C during the growth stage, especially under higher warming such as by end-of-century or under the higher warming scenario SSP585. This modeling approach can be applied to other almond growing regions and and other perennial crops such as other nut and fruit trees, after identifying appropriate ACIs.

The second study addresses limitations that often hinder regional applications of process-based crop models including flexibility of implementation, scalability of calibration, and representativeness across diverse climates, soils, and management. We develop a flexible, Python-based implementation of AquaCrop and propose a county-level calibration approach for U.S. maize. Relative to a field-scale calibration, our approach substantially improves AquaCrop performance with the U.S. production-

weighted average absolute bias decreasing from 28% to 16%, the SD error from 127% to 36%, the NRMSE from 47% to 26%, and the temporal correlation increasing from 0.32 to 0.42. Notably, these gains are achieved with calibration data from only 9 rainfed and 8 irrigated counties, enabled by a Sobol's sensitivity-guided selection of 11 parameters. This strategy balances accuracy and computational cost, broadening regional representativeness while remaining practical. The study also clarifies where variability arises in regional assessments. In the global analysis, spatial variability dominates overall yield variability; however, temporal variability contributes more at the U.S. state-scale than in global studies, highlighting the importance of accurately capturing interannual fluctuations for applications such as forecasting and early warning. By embedding multi-year, multi-county observations directly in the calibration, our approach improves AquaCrop performance in capturing spatial patterns and temporal dynamics of maize yields.

Leveraging the county-level calibrated AquaCrop, the third study evaluates U.S. maize outcomes with climate projections from 13 CMIP6 GCMs under SSP126 and SSP585 scenarios. It finds a multi-model median decline of 20% in national maize production by end-of-century under SSP585 and a smaller 1% decline under SSP126. The Corn Belt shows severe yield reductions between 10–20%. To inform adaptation, this study quantifies production responses to irrigating all maize-growing croplands and applying optimal fertilization. Either measure alone offsets climate damages and generates gains at mid-century, and combined they increase national production by nearly 30%. Under SSP585, only the joint implementation can fully counteract negative impacts by the end of the century. The analysis also explores relocating current maize croplands to areas with higher yield potentials.

Finally, this dissertation provides practical and scalable tools the agricultural industry can use to adapt to climate change: a statistical modeling approach for perennial systems and a scalable calibrated process-based crop model for annual systems. Together, they help assess climate risk and investigate adaptation options, such as irrigation, fertilization, and strategic relocation, at decision-relevant (county and regional) scales. Both approaches can be applied to other crops and regions, supporting climate-resilient planning and management.

References

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, *33*(1), 121–131.

Abatzoglou, J. T., & Brown, T. J. (2012). A comparison of statistical downscaling methods suited for wildfire applications. *International journal of climatology*, *32*(5), 772–780.

Abedinpour, M. (2021). The comparison of dssat-ceres and aquacrop models for wheat under water–nitrogen interactions. *Communications in Soil Science and Plant Analysis*, *52*(17), 2002–2017.

Adaskaveg, J. E., Michailides, T., & Eskalen, A. (2022). Fungicides, bactericides, biocontrols, and natural products for deciduous tree fruit and nut, citrus, strawberry, and vine crops in California. *University of California, Davis*.

Allen, J. C. (1976). A modified sine wave method for calculating degree days. *Environmental Entomology*, *5*(3), 388–396.

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*(9), D05109.

Almond Board of California. (n.d.). *Almond Lifecycle.* `https://www.almonds.com/why-almonds/almond-lifecycle`. ([Accessed 14-Oct-2022])

Alston, J. M., Carman, H. F., Christian, J. E., Dorfman, J., Murua, J. R., & Sexton, R. J. (1995). Optimal Reserve and Export Policies for the California Almond Industry: Theory, Econometrics and Simulations.

Amirouche, M., Smadhi, D., & Zella, L. (2021). Calibration and validation of the aquacrop model for the culture lettuce (lactuca sativa l.) under fertilization levels in pluvial condition. *Agricultural Engineering International: CIGR Journal*, *23*(1).

Angulo, C., Rötter, R., Lock, R., Enders, A., Fronzek, S., & Ewert, F. (2013). Implication of crop model calibration strategies for assessing regional impacts of climate change in europe. *Agricultural and Forest Meteorology*, *170*, 32–46.

Askari, H., & Cummings, J. T. (1977). Estimating agricultural supply response with the Nerlove model: a survey. *International economic review*, 257–292.

Balba, M. A., & Bray, R. H. (1956). The application of the mitscherlich equation for the calculation of plant composition due to fertilizer increments. *Soil Science Society of America Journal*, *20*(4), 515–518.

Baldocchi, D., & Wong, S. (2008). Accumulated winter chill is decreasing in the fruit growing regions of California. *Climatic Change*, *87*, 153–166.

Batjes, N. H. (2006). Isric-wise derived soil properties on a 5 by 5 arc-minutes global grid (version 1.0). *Wageningen: ISRIC—World Soil Information*.

Battisti, R., Sentelhas, P. C., & Boote, K. J. (2017). Inter-comparison of performance of soybean crop simulation models and their ensemble in southern brazil. *Field Crops Research*, *200*, 28–37.

Benmoussa, H., Ghrab, M., Mimoun, M. B., & Luedeling, E. (2017). Chilling and heat requirements for local and foreign almond (*Prunus dulcis* (Mill.) cultivars in a warm Mediterranean location based on 30 years of phenology records. *Agricultural and Forest Meteorology*, *239*, 34–46.

Beyer, R. M., Hua, F., Martin, P. A., Manica, A., & Rademacher, T. (2022). Relocating croplands could drastically reduce the environmental impacts of global food production. *Communications Earth & Environment*, *3*(1), 49.

Blanc, E., Caron, J., Fant, C., & Monier, E. (2017). Is current irrigation sustainable in the United States? An integrated assessment of climate change impact on water resources and irrigated crop yields. *Earth's future*, *5*(8), 877–892.

Blann, K. L., Anderson, J. L., Sands, G. R., & Vondracek, B. (2009). Effects of agricultural drainage on aquatic ecosystems: a review. *Critical reviews in environmental science and technology*, *39*(11), 909–1001.

Bo, Y., Liang, H., Li, T., & Zhou, F. (2024). Process-based modeling framework for sustainable irrigation management at the regional scale: integrating rice production, water use, and greenhouse gas emissions. *Geoscientific Model Development Discussions*, *2024*, 1–29.

Boryan, C., Yang, Z., Mueller, R., & Craig, M. (2011). Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program. *Geocarto International*, *26*(5), 341–358.

Brittain, C., Kremen, C., & Klein, A.-M. (2013). Biodiversity buffers pollination from changes in environmental conditions. *Global change biology*, *19*(2), 540–547.

Buchner, R., Duncan, R., Edstrom, J., Micke, W., Verdegaal, P., Viveros, M., & Connell, J. (2001). Evaluation of potential rootstocks for California almonds. In *Iii international symposium on pistachios and almonds 591* (pp. 45–52).

Busschaert, L., de Roos, S., Thiery, W., Raes, D., & De Lannoy, G. J. (2022). Net irrigation requirement under different climate scenarios using aquacrop over europe. *Hydrology and Earth System Sciences*, *26*(14), 3731–3752.

California Department of Food and Agriculture. (2022). *California Agricultural Statistics Review 2020–2021.* Retrieved from `https://www.cdfa.ca.gov/Statistics` ([Accessed 14-Aug-2022])

Campalans, A., Pagès, M., & Messeguer, R. (2001). Identification of differentially expressed genes by the cDNA-AFLP technique during dehydration of almond (*Prunus amygdalus*). *Tree Physiology*, *21*(10), 633–643.

Campoy, J. A., Ruiz, D., & Egea, J. (2011). Dormancy in temperate fruit trees in a global warming context: a review. *Scientia Horticulturae*, *130*(2), 357–372.

Caparas, M., Zobel, Z., Castanho, A. D., & Schwalm, C. R. (2021). Increasing risks of crop failure and water scarcity in global breadbaskets by 2030. *Environmental Research Letters*, *16*(10), 104013.

Carins-Murphy, M. R., Cochard, H., Deans, R. M., Gracie, A. J., & Brodribb, T. J. (2023). Combined heat and water stress leads to local xylem failure and tissue damage in pyrethrum flowers. *Plant Physiology*, *193*(1), 356–370.

Caswell, M., & Zilberman, D. (1985). The choices of irrigation technologies in california. *American journal of agricultural economics*, *67*(2), 224–234.

Cerrato, M., & Blackmer, A. (1990). Comparison of models for describing; corn yield response to nitrogen fertilizer. *Agronomy journal*, *82*(1), 138–143.

Chapman, S. C. (2008). Use of crop models to understand genotype by environment interactions for drought in real-world and simulated plant breeding trials. *Euphytica*, *161*(1), 195–208.

Chemeris, A., Liu, Y., & Ker, A. P. (2022). Insurance subsidies, climate change, and innovation: Implications for crop yield resiliency. *Food Policy*, *108*, 102232.

Chen, X., Bai, H., Xue, Q., Zhao, J., Zhao, C., & Feng, L. (2023). Wheatsm v5. 0: A python-based wheat growth and development simulation model with cloud services integration to enhance agricultural applications. *Agronomy*, *13*(9), 2411.

Chong, I.-G., & Jun, C.-H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and intelligent laboratory systems*, *78*(1-2), 103–112.

Chun, J. A., Li, S., Wang, Q., Lee, W.-S., Lee, E.-J., Horstmann, N., ... others (2016). Assessing rice productivity and adaptation strategies for southeast asia under climate change through multi-scale crop modeling. *Agricultural Systems*, *143*, 14–21.

Connell, J. H. (2000). Pollination of almonds: practices and problems. *HortTechnology*, *10*(1), 116–119.

Cook, R. D. (2000). Detection of influential observation in linear regression. *Technometrics*, *42*(1), 65–68.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., ... Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous united states. *International Journal of Climatology: a Journal of the Royal Meteorological Society*, *28*(15), 2031–2064.

Deppermann, A., Balkovič, J., Bundle, S.-C., Di Fulvio, F., Havlik, P., Leclère, D., ... Schepaschenko, D. (2018). Increasing crop production in russia and ukraine—regional and global impacts from intensification and recultivation. *Environmental Research Letters*, *13*(2), 025008.

de Roos, S., De Lannoy, G. J., & Raes, D. (2021). Performance analysis of regional aquacrop (v6. 1) biomass and surface soil moisture simulations using satellite and in situ observations. *Geoscientific Model Development*, *14*(12), 7309–7328.

Deryng, D., Conway, D., Ramankutty, N., Price, J., & Warren, R. (2014). Global crop yield response to extreme heat stress under multiple climate change futures. *Environmental Research Letters*, *9*(3), 034011.

Dhanoa, M. S., Sanderson, R., Cardenas, L., Shepherd, A., Chadwick, D., Powell, C., ... France, J. (2022). Overview and application of the mitscherlich equation and its extensions to estimate the soil nitrogen pool fraction associated with crop yield and nitrous oxide emission. *Advances in Agronomy*, *174*, 269–295.

Dieter, C., Linsey, K., Caldwell, R., Harris, M., Ivahnenko, T., Lovelace, J., ... Barber, N. (2018). Estimated use of water in the united states county-level data for 2015 (ver. 2.0, june 2018). *US Geological Survey data release*, *10*.

Djalovic, I., Kundu, S., Bahuguna, R. N., Pareek, A., Raza, A., Singla-Pareek, S. L., ... Varshney, R. K. (2024). Maize and heat stress: Physiological, genetic, and molecular insights. *The plant genome*, *17*(1), e20378.

Dlamini, L., Crespo, O., van Dam, J., & Kooistra, L. (2023). A global systematic review of improving crop model estimations by assimilating remote sensing data: implications for small-scale agricultural systems. *Remote Sensing*, *15*(16), 4066.

Driedonks, N., Rieu, I., & Vriezen, W. H. (2016). Breeding for plant heat tolerance at vegetative and reproductive stages. *Plant reproduction*, *29*, 67–79.

Duan, Q., Gupta, V. K., & Sorooshian, S. (1993). Shuffled complex evolution approach for effective and efficient global minimization. *Journal of optimization theory and applications*, *76*, 501–521.

Duan, Q., Sorooshian, S., & Gupta, V. K. (1994). Optimal use of the sce-ua global optimization method for calibrating watershed models. *Journal of hydrology*, *158*(3-4), 265–284.

Elias, E., Schrader, T. S., Abatzoglou, J. T., James, D., Crimmins, M., Weiss, J., & Rango, A. (2018). County-level climate change information to support decision-making on working lands. *Climatic Change*, *148*(3), 355–369.

Elliott, J., Müller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K., Büchner, M., ... others (2015). The global gridded crop model intercomparison: data and modeling protocols for phase 1 (v1. 0). *Geoscientific Model Development*, *8*(2), 261–277.

Epstein, L., Bassein, S., Zalom, F., & Wilhoit, L. (2001). Changes in pest management practice in almond orchards during the rainy season in California, USA. *Agriculture, ecosystems & environment*, *83*(1-2), 111–120.

Erenstein, O., Jaleta, M., Sonder, K., Mottaleb, K., & Prasanna, B. M. (2022). Global maize production, consumption and trade: trends and r&d implications. *Food security*, *14*(5), 1295–1319.

Er-Raki, S., Bouras, E., Rodriguez, J., Watts, C., Lizarraga-Celaya, C., & Chehbouni, A. (2021). Parameterization of the aquacrop model for simulating table grapes growth and water productivity in an arid region of mexico. *Agricultural Water Management*, *245*, 106585.

Ewert, F., Rounsevell, M., Reginster, I., Metzger, M., & Leemans, R. (2005). Future scenarios of European agricultural land use: I. Estimating changes in crop productivity. *Agriculture, Ecosystems & Environment*, *107*(2-3), 101–116.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958.

FAO. (2024). FAOSTAT Statistical Database.

FAO. (2024). The state of food security and nutrition in the world 2024.

Fazel, F., Ansari, H., & Aguilar, J. (2023). Determination of the most efficient forage sorghum irrigation scheduling strategies in the us central high plains using the aquacrop model and field experiments. *Agronomy*, *13*(10), 2446.

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

Feleke, H. G., Savage, M., & Tesfaye, K. (2021). Calibration and validation of apsim–maize, dssat ceres–maize and aquacrop models for ethiopian tropical environments. *South African Journal of Plant and Soil*, *38*(1), 36–51.

Fischer, G., Nachtergaele, F., Prieler, S., Van Velthuizen, H., Verelst, L., & Wiberg, D. (2008). Global agro-ecological zones assessment for agriculture (gaez 2008). *IIASA, Laxenburg, Austria and FAO, Rome, Italy*, *10*.

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., ... others (2011). Solutions for a cultivated planet. *Nature*, *478*(7369), 337–342.

Franke, J. A., Müller, C., Elliott, J., Ruane, A. C., Jägermeyr, J., Balkovic, J., ... others (2020). The ggcmi phase 2 experiment: global gridded crop model simulations under uniform changes in co 2, temperature, water, and nitrogen levels (protocol version 1.0). *Geoscientific Model Development*, *13*(5), 2315–2336.

Freitas, T. R., Santos, J. A., Silva, A. P., Fonseca, A., & Fraga, H. (2023). Evaluation of historical and future thermal conditions for almond trees in north-eastern Portugal. *Climatic Change*, *176*(7), 89.

García-Lara, S., & Serna-Saldivar, S. O. (2019). Corn history and culture. *Corn*, 1–18.

García-Vila, M., & Fereres, E. (2012). Combining the simulation crop model aquacrop with an economic model for the optimization of irrigation management at farm level. *European Journal of Agronomy*, *36*(1), 21–31.

Gradziel, T. M. (2009). Almond (*Prunus dulcis*) breeding. *Breeding plantation tree crops: temperate species*, 1–31.

Gradziel, T. M. (2022). Transfer of self-fruitfulness to cultivated almond from peach and wild almond. *Horticulturae*, *8*(10), 965.

Gradziel, T. M., Lampinen, B., Connell, J. H., & Viveros, M. (2007). Winters' almond: an early-blooming, productive, and high-quality pollenizer for 'Nonpareil. *HortScience*, *42*(7), 1725–1727.

Gradziel, T. M., & Martínez-Gómez, P. (2013). Almond breeding. *Plant breeding reviews*, *37*, 207–258.

Grassini, P., van Bussel, L. G., Van Wart, J., Wolf, J., Claessens, L., Yang, H., ... Cassman, K. G. (2015). How good is good enough? data requirements for reliable crop yield simulations and yield-gap analysis. *Field Crops Research*, *177*, 49–63.

Gregory, P. J., Ingram, J. S., & Brklacich, M. (2005). Climate change and food security. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1463), 2139–2148.

Guillamón, J. G., Egea, J., Mañas, F., Egea, J. A., & Dicenta, F. (2022). Risk of extreme early frosts in almond. *Horticulturae*, *8*(8), 687.

Hammer, G., Cooper, M., Tardieu, F., Welch, S., Walsh, B., van Eeuwijk, F., ... Podlich, D. (2006). Models for navigating biological complexity in breeding improved crop plants. *Trends in plant science*, *11*(12), 587–593.

Haqiqi, I., Grogan, D. S., Hertel, T. W., & Schlenker, W. (2021). Quantifying the impacts of compound extremes on agriculture. *Hydrology and Earth System Sciences*, *25*(2), 551–564.

Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, *90*(8), 1095–1108.

Hawkins, E., & Sutton, R. (2011). The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics*, *37*(1), 407–418. doi: 10.1007/s00382-010-0810-6

Haworth, M., Marino, G., Brunetti, C., Killi, D., De Carlo, A., & Centritto, M. (2018). The impact of heat stress and water deficit on the photosynthetic and stomatal physiology of olive (Olea europaea L.)—A case study of the 2017 heat wave. *Plants*, *7*(4), 76.

Hayhoe, K., Cayan, D., Field, C. B., Frumhoff, P. C., Maurer, E. P., Miller, N. L., ... Verville, J. H. (2004). Emissions pathways, climate change, and impacts on California. *Proceedings of the National Academy of Sciences*, *101*(34), 12422–12427. doi: 10.1073/pnas.0404500101

Hayman, P., & Thomas, D. (2017). A calendar to link almond phenology to climate risk. In *Vii international symposium on almonds and pistachios 1219* (pp. 137–142).

Hedhly, A. (2011). Sensitivity of flowering plant gametophytes to temperature fluctuations. *Environmental and Experimental Botany*, *74*, 9–16.

Henselek, Y., Eilers, E. J., Kremen, C., Hendrix, S. D., & Klein, A.-M. (2018). Pollination requirements of almond (*Prunus dulcis*): combining laboratory and field experiments. *Journal of economic entomology*, *111*(3), 1006–1013.

Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., ... others (2014). Apsim–evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software*, *62*, 327–350.

Hong, C., Mueller, N. D., Burney, J. A., Zhang, Y., AghaKouchak, A., Moore, F. C., ... Davis, S. J. (2020). Impacts of ozone and climate change on yields of perennial crops in california. *Nature Food*, *1*(3), 166–172.

Houska, T., Kraft, P., Chamorro-Chavez, A., & Breuer, L. (2015). Spotting model parameters using a ready-made python package. *PloS one*, *10*(12), e0145180.

Hsiao, T. C., Heng, L., Steduto, P., Rojas-Lara, B., Raes, D., & Fereres, E. (2009). Aquacrop—the fao crop model to simulate yield response to water: Iii. parameterization and testing for maize. *Agronomy Journal*, *101*(3), 448–459.

Hultgren, A., Carleton, T., Delgado, M., Gergel, D. R., Greenstone, M., Houser, T., ... others (2025). Impacts of climate change on global agriculture accounting for adaptation. *Nature*, *642*(8068), 644–652.

Iizumi, T., Shiogama, H., Imada, Y., Hanasaki, N., Takikawa, H., & Nishimori, M. (2018). Crop production losses associated with anthropogenic climate change. *International Journal of Climatology*, *38*, e1–e13. doi: 10.1002/joc.5389

Iizumi, T., Yokozawa, M., Sakurai, G., Travasso, M. I., Romanenkov, V., Oettli, P., ... Furuya, J. (2014). Historical changes in global yields: major cereal and legume crops from 1982 to 2006. *Global ecology and biogeography*, *23*(3), 346–357.

IPCC. (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.* Intergovernmental Panel on Climate Change. Retrieved from `https://www.ipcc.ch/srccl/`

Irmak, S., Sandhu, R., & Kukal, M. (2022). Multi-model projections of trade-offs between irrigated and rainfed maize yields under changing climate and future emission scenarios. *Agricultural Water Management*, *261*, 107344.

Jägermeyr, J., Müller, C., Ruane, A. C., Elliott, J., Balkovic, J., Castillo, O., ... others (2021). Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nature Food*, *2*(11), 873–885.

Jahanzad, E., Holtz, B. A., Zuber, C. A., Doll, D., Brewer, K. M., Hogan, S., & Gaudin, A. C. (2020). Orchard recycling improves climate change adaptation and mitigation potential of almond production systems. *PLoS One*, *15*(3), e0229588.

Jang, W., Neff, J., Im, Y., Doro, L., & Herrick, J. (2021). The hidden costs of land degradation in us maize agriculture. *Earth's Future*, *9*(2), e2020EF001641.

Johnston, R. Z., Sandefur, H. N., Bandekar, P., Matlock, M. D., Haggard, B. E., & Thoma, G. (2015). Predicting changes in yield and water use in the production of corn in the united states under climate change scenarios. *Ecological Engineering*, *82*, 555–565.

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., ... Ritchie, J. T. (2003). The dssat cropping system model. *European journal of agronomy*, *18*(3-4), 235–265.

Kakoulidou, I., Avramidou, E. V., Baránek, M., Brunel-Muguet, S., Farrona, S., Johannes, F., ... Maury, S. (2021). Epigenetics for Crop Improvement in Times of Global Change. *Biology*, *10*(8). doi: 10.3390/biology10080766

Kang, Y., Khan, S., & Ma, X. (2009). Climate change impacts on crop yield, crop water productivity and food security–a review. *Progress in natural Science*, *19*(12), 1665–1674.

Keating, B. A., Carberry, P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., ... others (2003). An overview of apsim, a model designed for farming systems simulation. *European journal of agronomy*, *18*(3-4), 267–288.

Kebede, E. A., Oluoch, K. O., Siebert, S., Mehta, P., Hartman, S., Jägermeyr, J., ... others (2025). A global open-source dataset of monthly irrigated and rainfed cropped areas (mirca-os) for the 21st century. *Scientific Data*, *12*(1), 208.

Kelly, T., & Foster, T. (2021). Aquacrop-ospy: Bridging the gap between research and practice in crop-water modeling. *Agricultural Water Management*, *254*, 106976.

Khasraei, A., & Eslamian, S. (2024). Investigating the effects of agricultural management on wheat plant productivity gap using aqua-crop software (case study: Tabriz city). In *Handbook of climate change impacts on river basin management* (pp. 90–108). CRC Press.

Kim, D., & Kaluarachchi, J. (2015). Validating fao aquacrop using landsat images and regional crop information. *Agricultural Water Management*, *149*, 143–155.

Kim, K.-H., & Lee, B.-M. (2023). Effects of climate change and drought tolerance on maize growth. *Plants*, *12*(20), 3548.

Kodad, O., Socias, R., & i COMPANY. (2009). Effect of pollination time on fruit set in an autogamous almond cultivar. *The Journal of Horticultural Science and Biotechnology*, *84*(3), 350–354.

Kraft, P., Multsch, S., Vaché, K., Frede, H.-G., & Breuer, L. (2010). Using python as a coupling platform for integrated catchment models. *Advances in Geosciences*, *27*, 51–56.

Kreitzman, M., Toensmeier, E., Chan, K. M., Smukler, S., & Ramankutty, N. (2020). Perennial staple crops: Yields, distribution, and nutrition in the global food system. *Frontiers in Sustainable Food Systems*, *4*, 588988.

Kumar, C., Dhillon, J., Huang, Y., & Reddy, K. (2025). Explainable machine learning models for corn yield prediction using uav multispectral data. *Computers and Electronics in Agriculture*, *231*, 109990.

Kumar, M. (2016). Impact of climate change on crop yield and role of model for achieving food security. *Environmental Monitoring and Assessment*, *188*(8), 465.

Lafferty, D. C., & Sriver, R. L. (2023). Downscaling and bias-correction contribute considerable uncertainty to local climate projections in CMIP6. *npj Climate and Atmospheric Science*, *6*(1), 158.

Lamsal, G., & Marston, L. T. (2025). Monthly crop water consumption of irrigated crops in the united states from 1981 to 2019. *Water Resources Research*, *61*(2), e2024WR038334.

Lan, K., Chen, X., Ridoutt, B. G., Huang, J., & Scherer, L. (2021). Closing yield and harvest area gaps to mitigate water scarcity related to china's rice production. *Agricultural Water Management*, *245*, 106602.

Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., ... Hawkins, E. (2020). Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth System Dynamics*, *11*(2), 491–508.

Leisner, C. P. (2020). Climate change impacts on food security-focus on perennial cropping systems and nutritional value. *Plant Science*, *293*, 110412.

Lesk, C., Rowhani, P., & Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature*, *529*(7584), 84–87. doi: 10.1038/nature16467

Li, C., Camac, J., Robinson, A., & Kompas, T. (2025). Predicting changes in agricultural yields under climate change scenarios and their implications for global food security. *Scientific Reports*, *15*(1), 2858.

Li, F., Yu, D., & Zhao, Y. (2019). Irrigation scheduling optimization for cotton based on the aquacrop model. *Water resources management*, *33*(1), 39–55.

Livneh, B., Bohn, T. J., Pierce, D. W., Munoz-Arriola, F., Nijssen, B., Vose, R., ... Brekke, L. (2015). A spatially comprehensive, hydrometeorological data set for mexico, the us, and southern canada 1950–2013. *Scientific data*, *2*(1), 1–12.

Lobell, D. B., Cahill, K. N., & Field, C. B. (2007). Historical effects of temperature and precipitation on California crop yields. *Climatic change*, *81*(2), 187–203.

Lobell, D. B., Deines, J. M., & Tommaso, S. D. (2020). Changes in the drought sensitivity of us maize yields. *Nature Food*, *1*(11), 729–735.

Lobell, D. B., & Field, C. B. (2011). California perennial crops in a changing climate. *Climatic Change*, *109*, 317–333.

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*(2-4), 208–218.

Lobell, D. B., & Gourdji, S. M. (2012). The influence of climate change on global crop productivity. *Plant physiology*, *160*(4), 1686–1697.

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*(5), 497–501.

Lobell, D. B., Roberts, M. J., Schlenker, W., Braun, N., Little, B. B., Rejesus, R. M., & Hammer, G. L. (2014). Greater sensitivity to drought accompanies maize yield increase in the us midwest. *Science*, *344*(6183), 516–519.

Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. *Science*, *333*(6042), 616–620. doi: 10.1126/science.1204531

Lorite, I. J., García-Vila, M., Santos, C., Ruiz-Ramos, M., & Fereres, E. (2013). Aquadata and aquagis: two computer utilities for temporal and spatial simulations of water-limited yield with aquacrop. *Computers and Electronics in Agriculture*, *96*, 227–237.

Luedeling, E., Zhang, M., & Girvetz, E. H. (2009). Climatic changes lead to declining winter chill for fruit and nut trees in California during 1950–2099. *PloS one*, *4*(7), e6166.

Maestrini, B., & Basso, B. (2018). Drivers of within-field spatial and temporal variability of crop yield across the us midwest. *Scientific reports*, *8*(1), 14833.

Maestrini, B., Mimić, G., van Oort, P. A., Jindo, K., Brdar, S., van Evert, F. K., & Athanasiados, I. (2022). Mixing process-based and data-driven approaches in yield prediction. *European Journal of Agronomy*, *139*, 126569.

Martínez-García, P. J., Kodad, O., Gouta, H., Devin, S. R., Prudencio, A. S., Rubio, M., & Martínez-Gómez, P. (2022). Genomic Designing for Drought Tolerant Almond Varieties. , 161–175.

Martínez-Gómez, P., Prudencio, A. S., Gradziel, T. M., & Dicenta, F. (2017). The delay of flowering time in almond: a review of the combined effect of adaptation, mutation and breeding. *Euphytica*, *213*, 1–10.

Medellín-Azuara, J., Howitt, R. E., MacEwan, D. J., & Lund, J. R. (2011). Economic impacts of climate-related changes to California agriculture. *Climatic Change*, *109*, 387–405.

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., ... Bailey, A. A. (2004). The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research: Atmospheres*, *109*(D7). doi: https://doi.org/10.1029/2003JD003823

Molotoks, A., Smith, P., & Dawson, T. P. (2021). Impacts of land use, population, and climate change on global food security. *Food and energy security*, *10*(1), e261.

Monfreda, C., Ramankutty, N., & Foley, J. A. (2008). Farming the planet: 2. geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global biogeochemical cycles*, *22*(1).

Monier, E., Gao, X., Scott, J. R., Sokolov, A. P., & Schlosser, C. A. (2015). A framework for modeling uncertainty in regional climate change. *Climatic Change*, *131*, 51–66.

Monier, E., Xu, L., & Snyder, R. (2016). Uncertainty in future agro-climate projections in the United States and benefits of greenhouse gas mitigation. *Environmental Research Letters*, *11*(5), 055001.

Moore, F. C., Baldos, U. L. C., & Hertel, T. (2017). Economic impacts of climate change on agriculture: a comparison of process-based and statistical yield models. *Environmental Research Letters*, *12*(6), 065008.

Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., & Foley, J. A. (2012). Closing yield gaps through nutrient and water management. *Nature*, *490*(7419), 254–257.

Muhammad, S., Sanden, B. L., Lampinen, B. D., Saa, S., Siddiqui, M. I., Smart, D. R., ... Brown, P. H. (2015). Seasonal changes in nutrient content and concentrations in a mature deciduous tree species: Studies in almond (*Prunus dulcis* (Mill.) D.A. Webb). *European Journal of Agronomy*, *65*, 52–68.

Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais, P., ... others (2017). Global gridded crop model evaluation: benchmarking, skills, deficiencies and implications. *Geoscientific Model Development*, *10*(4), 1403–1422.

Müller, C., Elliott, J., Kelly, D., Arneth, A., Balkovic, J., Ciais, P., ... others (2019). The global gridded crop model intercomparison phase 1 simulation dataset. *Scientific data*, *6*(1), 50.

Myers, S. S., Smith, M. R., Guth, S., Golden, C. D., Vaitla, B., Mueller, N. D., ... Huybers, P. (2017). Climate change and global food systems: potential impacts on food security and undernutrition. *Annual Review of Public Health*, *38*, 259–277. doi: 10.1146/annurev-publhealth-031816-044356

National Academies of Sciences, E., Medicine, et al. (2023). *Toward a 21st century national data infrastructure: Enhancing survey programs by using multiple data sources*.

Nguyen, H., Thompson, A., & Costello, C. (2023). Impacts of historical droughts on maize and soybean production in the southeastern united states. *Agricultural Water Management*, *281*(108237), 1336–1347.

Olba-Zięty, E., Stolarski, M. J., & Krzyżaniak, M. (2021). Economic evaluation of the production of perennial crops for energy purposes—a review. *Energies*, *14*(21), 7147.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., ... Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. doi: 10.5194/gmd-9-3461-2016

Ortega, E., Egea, J., & Dicenta, F. (2004). Effective pollination period in almond cultivars. *HortScience*, *39*(1), 19–22.

Palacio-Bielsa, A., Cambra, M., Martínez, C., Olmos, A., Pallás, V., López, M. M., ... Esmenjaud, D. (2017). Almond diseases. *CABI*, 321–374. doi: 10.1079/9781780643540.0321

Pardey, P. G., Alston, J. M., & Ruttan, V. W. (2010). The economics of innovation and technical change in agriculture. *Handbook of the Economics of Innovation*, *2*, 939–984.

Parker, L. E., & Abatzoglou, J. T. (2018). Shifts in the thermal niche of almond under climate change. *Climatic change*, *147*(1-2), 211–224.

Parker, L. E., McElrone, A. J., Ostoja, S. M., & Forrestel, E. J. (2020). Extreme heat effects on perennial crops and strategies for sustaining future production. *Plant Science*, *295*, 110397.

Partridge, T., Winter, J., Kendall, A., Basso, B., Pei, L., & Hyndman, D. (2023). Irrigation benefits outweigh costs in more us croplands by mid-century. *Communications Earth & Environment*, *4*(1), 274.

Pasquel, D., Roux, S., Richetti, J., Cammarano, D., Tisseyre, B., & Taylor, J. A. (2022). A review of methods to evaluate crop model performance at multiple and changing spatial scales. *Precision Agriculture*, *23*(4), 1489–1513.

Paudel, D., De Wit, A., Boogaard, H., Marcos, D., Osinga, S., & Athanasiadis, I. N. (2023). Interpretability of deep learning models for crop yield forecasting. *Computers and Electronics in Agriculture*, *206*, 107663.

Pierce, D. W., Cayan, D. R., Feldman, D. R., & Risser, M. D. (2023). Future increases in North American Extreme Precipitation in CMIP6 downscaled with LOCA. *Journal of Hydrometeorology*, *24*(5), 951–975.

Pierce, D. W., Cayan, D. R., & Thrasher, B. L. (2014). Statistical downscaling using localized constructed analogs (LOCA). *Journal of Hydrometeorology*, *15*(6), 2558–2585.

Pope, K., Da Silva, D., Brown, P., & DeJong, T. (2014). A biologically based approach to modeling spring phenology in temperate deciduous trees. *Agricultural and Forest Meteorology*, *198*, 15–23.

Pope, K. S., Dose, V., Da Silva, D., Brown, P. H., & DeJong, T. M. (2015). Nut crop yield records show that budbreak-based chilling requirements may not reflect yield decline chill thresholds. *International journal of biometeorology*, *59*, 707–715.

Portmann, F. T., Siebert, S., & Döll, P. (2010). Mirca2000—global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global biogeochemical cycles*, *24*(1).

Portmann, R. W., Solomon, S., & Hegerl, G. C. (2009). Spatial and seasonal patterns in climate change, temperatures, and precipitation across the united states. *Proceedings of the National Academy of Sciences*, *106*(18), 7324–7329.

Potter, P., Ramankutty, N., Bennett, E. M., & Donner, S. D. (2010). Characterizing the spatial patterns of global fertilizer application and manure production. *Earth interactions*, *14*(2), 1–22.

Raes, D., Steduto, P., Hsiao, T., & Fereres, E. (2023a). Aquacrop reference manual: Annexes. *Reference Manual: AquaCrop, Version 7.1*.

Raes, D., Steduto, P., Hsiao, T., & Fereres, E. (2023b). Aquacrop reference manual: Chapter 3 calculation procedures. *Reference Manual: AquaCrop, Version 7.1*.

Raes, D., Steduto, P., Hsiao, T. C., & Fereres, E. (2009). Aquacrop—the fao crop model to simulate yield response to water: Ii. main algorithms and software description. *Agronomy Journal*, *101*(3), 438–447.

Ramírez, F., Kallarackal, J., Ramírez, F., & Kallarackal, J. (2015). The effect of increasing temperature on phenology. *Responses of fruit trees to global climate change*, 11–13.

Ramirez-Villegas, J., Koehler, A.-K., & Challinor, A. J. (2017). Assessing uncertainty and complexity in regional-scale crop model simulations. *European Journal of Agronomy*, *88*, 84–95.

Rattigan, K., & Hill, S. (1986). Relationship between temperature and flowering in almond. *Australian Journal of Experimental Agriculture*, *26*(3), 399–404.

Ray, D. K., Gerber, J. S., MacDonald, G. K., & West, P. C. (2015). Climate variation explains a third of global crop yield variability. *Nature Communications*, *6*, 5989. doi: 10.1038/ncomms6989

Reed, P. M., Hadjimichael, A., Moss, R. H., Brelsford, C., Burleyson, C. D., Cohen, S., ... Yoon, J. (2022). Multisector Dynamics: Advancing the Science of Complex Adaptive Human-Earth Systems. *Earth's Future*, *10*(3). doi: https://doi.org/10.1029/2021EF002621

Reed, P. M., Hadjimichael, A., Moss, R. H., Monier, E., Alba, S., Brelsford, C., ... Yoon, J. (2022). *MultiSector Dynamics: Scientific Challenges and a Research Vision for 2030, A Community of Practice Supported by the United States Department of Energy's Office of Science.* Zenodo. doi: 10.5281/zenodo.6144309

Ren, C., He, L., & Rosa, L. (2025). Integrated irrigation and nitrogen optimization is a resource-efficient adaptation strategy for us maize and soybean production. *Nature Food*, *6*(4), 389–400.

Reynolds, M., Kropff, M., Crossa, J., Koo, J., Kruseman, G., Molero Milan, A., ... others (2018). Role of modelling in international crop research: overview and some case studies. *Agronomy*, *8*(12), 291.

Rezaei, E. E., Webber, H., Asseng, S., Boote, K., Durand, J. L., Ewert, F., ... MacCarthy, D. S. (2023). Climate change impacts on crop yields. *nature reviews earth & environment*, *4*(12), 831–846.

Riahi, K., Vuuren, D. P. v., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., ... Tavoni, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, *42*, 153–168. doi: https://doi.org/10.1016/j.gloenvcha.2016.05.009

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., ... others (2011). Merra: Nasa's modern-era retrospective analysis for research and applications. *Journal of climate*, *24*(14), 3624–3648.

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*(9), 095010.

Rodríguez, A., Pérez-López, D., Sánchez, E., Centeno, A., Gómara, I., Dosio, A., & Ruiz-Ramos, M. (2019). Chilling accumulation in fruit trees in Spain under climate change. *Natural Hazards and Earth System Sciences*, *19*(5), 1087–1103.

Romero, P., Botia, P., & Garcia, F. (2004). Effects of regulated deficit irrigation under subsurface drip irrigation conditions on vegetative development and yield of mature almond trees. *Plant and Soil*, *260*(1-2), 169–181.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A. C., Müller, C., Arneth, A., ... others (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*(9), 3268–3273.

Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P., ... others (2013). The agricultural model intercomparison and improvement project (agmip): protocols and pilot studies. *Agricultural and forest meteorology*, *170*, 166–182.

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

Ruane, A. C., Phillips, M., Müller, C., Elliott, J., Jägermeyr, J., Arneth, A., ... others (2021). Strong regional influence of climatic forcing datasets on global crop model ensembles. *Agricultural and Forest Meteorology*, *300*, 108313.

Ruess, P., Konar, M., Wanders, N., & Bierkens, M. (2023). Irrigation by crop in the continental united states from 2008 to 2020. *Water Resources Research*, *59*(2), e2022WR032804.

Sacks, W. J., Deryng, D., Foley, J. A., & Ramankutty, N. (2010). Crop planting dates: an analysis of global patterns. *Global ecology and biogeography*, *19*(5), 607–620.

Sánchez, B., Rasmussen, A., & Porter, J. R. (2014). Temperatures and the growth and development of maize and rice: a review. *Global change biology*, *20*(2), 408–417.

Sandhu, R., & Irmak, S. (2019). Performance of aquacrop model in simulating maize growth, yield, and evapotranspiration under rainfed, limited and full irrigation. *Agricultural Water Management*, *223*, 105687.

Sassenrath, G. F., Heilman, P., Luschei, E., Bennett, G. L., Fitzgerald, G., Klesius, P., ... Zimba, P. V. (2008). Technology, complexity and change in agricultural production systems. *Renewable Agriculture and Food Systems*, *23*(4), 285–295.

Sauer, T., Havlík, P., Schneider, U. A., Schmid, E., Kindermann, G., & Obersteiner, M. (2010). Agriculture and resource availability in a changing world: The role of irrigation. *Water Resources Research*, *46*(6).

Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to us crop yields under climate change. *Proceedings of the National Academy of sciences*, *106*(37), 15594–15598.

Segura, J. M. A., i Company, R. S., & Kodad, O. (2017). Late-blooming in almond: A controversial objective. *Scientia Horticulturae*, *224*, 61–67.

Seidel, S. J., Palosuo, T., Thorburn, P., & Wallach, D. (2018). Towards improved calibration of crop models–where are we now and where should we go? *European Journal of Agronomy*, *94*, 25–35.

Sharafkhane, M. G., Ziaei, A. N., Naghedifar, S. M., Akbari, A., & Verdi, A. (2024). Aquacrop plug-in-pso: A novel irrigation scheduling optimization framework for maize to maximize crop water productivity using in-season weather forecast and crop yield estimation. *Agricultural Water Management*, *306*, 109153.

Sheffield, J., Goteti, G., & Wood, E. F. (2006). Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of climate*, *19*(13), 3088–3111.

Shi, T., Cao, J., Cao, J., Zhu, F., Cao, F., & Su, E. (2023). Almond (Amygdalus communis L.) Kernel Protein: A Review on the Extraction, Functional Properties and Nutritional Value. *Food Research International*, 112721.

Singh, R., Srivastava, P., Singh, P., Upadhyay, S., & Raghubanshi, A. S. (2017). Human overpopulation and food security: challenges for the agriculture sustainability. In *Environmental issues surrounding human overpopulation* (pp. 12–39). IGI Global.

Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and computers in simulation*, *55*(1-3), 271–280.

Soil Survey Staff. (2024). *Gridded Soil Survey Geographic (gSSURGO) Database for the Conterminous United States [DataSet].* `https://gdg.sc.egov.usda.gov/`. United States Department of Agriculture, Natural Resources Conservation Service. (Accessed: 20-Dec-2024)

Sparrow, R., & Howard, M. (2021). Robots in agriculture: prospects, impacts, ethics, and policy. *Precision Agriculture*, *22*(3), 818–833. doi: 10.1007/s11119-020-09757-9

Spinelli, G. M., Snyder, R. L., Sanden, B. L., Gilbert, M., & Shackel, K. A. (2018). Low and variable atmospheric coupling in irrigated Almond (*Prunus dulcis*) canopies indicates a limited influence of stomata on orchard evapotranspiration. *Agricultural Water Management*, *196*, 57–65.

Steduto, P., Hsiao, T. C., Raes, D., & Fereres, E. (2009). Aquacrop—the fao crop model to simulate yield response to water: I. concepts and underlying principles. *Agronomy Journal*, *101*(3), 426–437.

Sunding, D., & Zilberman, D. (2001). The agricultural innovation process: research and technology adoption in a changing agricultural sector. *Handbook of agricultural economics*, *1*, 207–261.

Tang, P., Li, N., Li, M., Zhang, F., Fu, Q., Xu, Y., & Liu, D. (2023). Rice irrigation water efficiency improvement: An aquacrop-based optimization modeling approach. *European Journal of Agronomy*, *148*, 126867.

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*(4), 485–498.

Tebaldi, C., Debeire, K., Eyring, V., Fischer, E., Fyfe, J., Friedlingstein, P., ... others (2021). Climate model projections from the scenario model intercomparison project (scenariomip) of cmip6. *Earth System Dynamics*, *12*(1), 253–293.

Thomas, D. (2019). Managing Almond Production in a Variable and Changing Climate; Hort Innovation: Sydney, Australia, 2019.

Thrasher, B., Wang, W., Michaelis, A., Melton, F., Lee, T., & Nemani, R. (2022). Nasa global daily downscaled projections, cmip6. *Scientific data*, *9*(1), 262.

Tigchelaar, M., Battisti, D. S., Naylor, R. L., & Ray, D. K. (2018). Future warming increases probability of globally synchronized maize production shocks. *Proceedings of the National Academy of Sciences*, *115*(26), 6644–6649.

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the national academy of sciences*, *108*(50), 20260–20264.

Tita, D., Mahdi, K., Devkota, K. P., & Devkota, M. (2025). Climate change and agronomic management: Addressing wheat yield gaps and sustainability challenges in the mediterranean and mena regions. *Agricultural Systems*, *224*, 104242.

Tombesi, S., Scalia, R., Connell, J., Lampinen, B., & DeJong, T. (2010). Fruit development in almond is influenced by early spring temperatures in California. *The Journal of Horticultural Science and Biotechnology*, *85*(4), 317–322.

Traynor, J. (2017). A history of almond pollination in California. *Bee World*, *94*(3), 69–79.

Tubiello, F. N., Soussana, J.-F., & Howden, S. M. (2007). Crop and pasture response to climate change. *Proceedings of the National Academy of Sciences*, *104*(50), 19686–19690.

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*(4), 045003.

USDA NASS. (2023). *National Agricultural Statistics Service[Dataset].* `https://www.nass.usda.gov`. (Accessed: 15-May-2024)

USDA NASS. (2025). Crop Values 2024 Summary.

USDA National Agricultural Statistics Service. (2021). *2020 california almond acreage report.* Retrieved from `https://www.nass.usda.gov/Statistics_by_State/California/Publications/Specialty_and_Other_Releases/Almond/Acreage/202104_almac.pdf` (Accessed: 2025-01-23)

USDA National Agricultural Statistics Service. (2022). *2022 census of agriculture: California county-level data, volume 1, chapter 2.* Retrieved from `https://www.nass.usda.gov/Publications/AgCensus/2022/Full_Report/Volume_1,_Chapter_2_County_Level/California/st06_2_008_008.pdf` (Accessed: 2025-01-23)

Vadez, V., Grondin, A., Chenu, K., Henry, A., Laplaze, L., Millet, E. J., & Carminati, A. (2024). Crop traits and production under drought. *Nature Reviews Earth & Environment*, *5*(3), 211–225.

Van Gaelen, H., Tsegay, A., Delbecque, N., Shrestha, N., Garcia, M., Fajardo, H., ... others (2015). A semi-quantitative approach for modelling crop response to soil fertility: evaluation of the aquacrop procedure. *The Journal of Agricultural Science*, *153*(7), 1218–1233.

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., ... Rose, S. K. (2011). The representative concentration pathways: an overview. , *109*(1), 5. doi: 10.1007/s10584-011-0148-z

Wallach, D., Palosuo, T., Thorburn, P., Hochman, Z., Gourdain, E., Andrianasolo, F., ... others (2021). The chaos in calibrating crop models: Lessons learned from a multi-model calibration exercise. *Environmental Modelling & Software*, *145*, 105206.

Warziniack, T., & Brown, T. C. (2019). The importance of municipal and agricultural demands in future water shortages in the united states. *Environmental Research Letters*, *14*(8), 084036.

Weih, M., Adam, E., Vico, G., & Rubiales, D. (2022). Application of crop growth models to assist breeding for intercropping: Opportunities and challenges. *Frontiers in Plant Science*, *13*, 720486.

Wheeler, T., & Von Braun, J. (2013). Climate change impacts on global food security. *Science*, *341*(6145), 508–513.

Wieder, W. (2014). *Regridded harmonized world soil database v1. 2* (Tech. Rep.). ORNL Distributed Active Archive Center (DAAC).

Williamson, H. F., Brettschneider, J., Caccamo, M., Davey, R. P., Goble, C., Kersey, P. J., ... others (2023). Data management challenges for artificial intelligence in plant and agricultural research. *F1000Research*, *10*, 324.

Wing, I. S., De Cian, E., & Mistry, M. N. (2021). Global vulnerability of crop yields to climate change. *Journal of Environmental Economics and Management*, *109*, 102462. doi: https://doi.org/10.1016/j.jeem.2021.102462

Wing, I. S., Monier, E., Stern, A., & Mundra, A. (2015). US major crops' uncertain climate change risks and greenhouse gas mitigation benefits. *Environmental Research Letters*, *10*(11), 115002.

Wood, A. W., Leung, L. R., Sridhar, V., & Lettenmaier, D. P. (2004). Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic change*, *62*(1), 189–216.

Wu, S., & Monier, E. (2025). A gridded and python implementation of aquacrop calibrated with county-level yields: A case study of maize in us maize).

Wu, S., Zikalala, P. G., Alba, S., Jarvis-Shean, K. S., Kisekka, I., Segaran, M., ... Monier, E. (2025). Advancing the modeling of future climate and innovation impacts on perennial crops to support adaptation: A case study of california almonds. *Earth's Future*, *13*(4), e2024EF005033.

Xiong, W., Skalskỳ, R., Porter, C. H., Balkovič, J., Jones, J. W., & Yang, D. (2016). Calibration-induced uncertainty of the epic model to estimate climate change impact on global maize yield. *Journal of Advances in Modeling Earth Systems*, *8*(3), 1358–1375.

Xue, J., Huo, Z., & Kisekka, I. (2021). Assessing impacts of climate variability and changing cropping patterns on regional evapotranspiration, yield and water productivity in California's San Joaquin watershed. *Agricultural Water Management*, *250*, 106852.

Yang, M., & Wang, G. (2023). Heat stress to jeopardize crop production in the us corn belt based on downscaled cmip5 projections. *Agricultural Systems*, *211*, 103746.

Zhang, L., Bai, G., Evett, S. R., Colaizzi, P. D., Xue, Q., Marek, G., ... Lin, X. (2025). Increased irrigation could mitigate future warming-induced maize yield losses in the ogallala aquifer. *Communications Earth & Environment*, *6*(1), 483.

Zhang, X.-Y., Trame, M. N., Lesko, L. J., & Schmidt, S. (2015). Sobol sensitivity analysis: a tool to guide the development and evaluation of systems pharmacology models. *CPT: pharmacometrics & systems pharmacology*, *4*(2), 69–79.

Zhang, Z., Jin, Y., Chen, B., & Brown, P. (2019). California almond yield prediction at the orchard level with a machine learning approach. *Frontiers in plant science*, *10*, 809.

Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., ... others (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of sciences*, *114*(35), 9326–9331.

Zhu, J., Parker, A., Gou, F., Agnew, R., Yang, L., Greven, M., ... others (2021). Developing perennial fruit crop models in apsim next generation using grapevine as an example. *in silico Plants*, *3*(2), diab021.