THE UNIVERSITY OF CHICAGO

SPATIAL DISPARITIES IN CLIMATE IMPACTS ON AGRICULTURAL PRODUCTION: A GLOBAL ASSESSMENT OF SOME ADAPTATION STRATEGIES AND APPLIED NUTRIENT IMPACTS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF THE GEOPHYSICAL SCIENCES

BY
JAMES ALLEN FRANKE

CHICAGO, ILLINOIS
JUNE 2023

Copyright © 2023 by James Allen Franke

All Rights Reserved

TABLE OF CONTENTS

LIST OF FIGURES vii

LIST OF TABLES x

ACKNOWLEDGMENTS xi

STATEMENT OF ORIGINALITY AND COPYRIGHT ATTRIBUTION xii

ABSTRACT xiii

1 INTRODUCTION 1

2 THE GGCMI PHASE 2 EXPERIMENT: GLOBAL GRIDDED CROP MODEL SIMULATIONS UNDER UNIFORM CHANGES IN CO_2, TEMPERATURE, WATER, AND NITROGEN LEVELS 5

2.1 Introduction 6

2.2 Simulation Objectives and Protocol 9

2.2.1 Goals 9

2.2.2 Modeling Protocol 9

2.2.3 Harmonization Between Models 14

2.2.4 Output Data Products 18

2.3 Models Contributing 21

2.3.1 Results 24

2.3.2 Evaluation of Model Performance 24

2.3.3 Model Crop Yield Responses Under CTWN Forcing 26

2.4 Discussion and Conclusions 34

2.5 Data Availability 36

3 THE GGCMI PHASE 2 EMULATORS: GLOBAL GRIDDED CROP MODEL RESPONSES TO CHANGES IN CO_2, TEMPERATURE, WATER, AND NITROGEN 37
3.1 Introduction . . . 38
3.2 Training Dataset . . . 42
3.2.1 The GGCMI Phase 2 Dataset . . . 42
3.2.2 Climatological vs. Year-to-Year Responses . . . 44
3.3 Emulation . . . 47
3.3.1 Statistical Model . . . 48
3.3.2 Feature Importance and Reduced Statistical Model . . . 49
3.3.3 Model Fitting . . . 52
3.4 Emulator Evaluation . . . 53
3.4.1 Yield Response . . . 54
3.4.2 Emulator Performance Metrics . . . 57
3.4.3 Emulation of Realistic Climate Projections . . . 63
3.5 Emulator Results and Products . . . 65
3.6 Discussion and Conclusions . . . 70
3.7 Data Availability . . . 73

4 AGRICULTURAL BREADBASKETS SHIFT POLEWARD GIVEN ADAPTIVE FARMER BEHAVIOR UNDER CLIMATE CHANGE . . . 74
4.1 Introduction . . . 76
4.2 Materials and Methods . . . 78
4.2.1 Crop Model Simulations . . . 78
4.2.2 Breadbasket Analysis . . . 83
4.3 Results . . . 84
4.3.1 Expected Future Growing-Season Climate Changes . . . 84
4.3.2 Stasis in the No-Adaptation Case . . . 86

4.3.3 Growing Season Adaptation Allows Geographic Shifts 89
4.4 Discussion 96
4.5 Data Availability 99

5 NUTRIENT CONTROLS ON CLIMATE RESPONSE AND UNCERTAINTY FOR MAJOR GRAINS 101
5.1 Introduction 101
5.2 Methods 104
5.3 Results 109
5.4 Discussion and Conclusions 113
5.5 Data Availability 116

6 CONCLUSIONS 117

REFERENCES 119

APPENDIX 144
S1 Supplemental Materials for Chapter 2 144
S1.1 Cultivation Areas 144
S1.2 Reanalysis Climate Products 145
S1.3 Model Details 147
S1.4 Results 150
S2 Supplemental Materials for Chapter 3 161
S2.1 Sampling 161
S2.2 Variability Changes in Future Climate Projections 163
S2.3 Yield Response for A1 (Growing Season Adaptation) simulations . . . 166
S2.4 Normalized Error for Other Cases 167
S2.5 Emulation of Yields in a Realistic Climate Simulation at High Latitude 169

S2.6 Emulator Products 170
S2.7 Reduced Specification (23-term) Emulator Examples 173
S2.8 Yield Responses for Other Crops and Models 179
S2.9 Cross Validation Error for All Models 181
S3 Supplemental Materials for Chapter 4 187
S3.1 Additional LPJmL Simulations 187
S3.2 Breadbasket and Cultivation Areas 188
S3.3 Model Comparison and CMIP6 Climate Velocity 188
S3.4 Results 188
S4 Supplemental Materials for Chapter 5 215

LIST OF FIGURES

2.1 GGCMI Phase 2 simulation coverage and global cropland 10
2.2 Assessment of crop model performance in GGCMI Phase 2 25
2.3 Illustration of the spatial patterns of yields 27
2.4 Model probability densities for soybean yields 28
2.5 Illustration of the distribution of regional yield changes for maize 29
2.6 Illustration of the distribution of regional yield changes for soybeans and rice . . 30

3.1 Distinction between crop yield responses to year-to-year and mean shifts 45
3.2 Increased crop-yield sensitivity to year-to-year climate variations 47
3.3 Results from the polynomial feature selection process 51
3.4 Spatial pattern in baseline yield successfully captured by the emulator 55
3.5 Spatial variations in yield response successfully captured by the emulator 56
3.6 Variations in response across models successfully captured by the emulator . . . 56
3.7 Assessment of emulator performance over currently cultivated areas 58
3.8 Test of emulator performance 59
3.9 Test of emulator performance with a realistic climate projection 64
3.10 Emulated global damage functions 66
3.11 The factors affecting yields in a more realistic climate scenario 68

4.1 Poleward shift of growing season temperatures under warming 85
4.2 Yields decline while the latitude of peak yield remains fixed 86
4.3 Warming-related changes in zonal potential production without adaptation . . . 88
4.4 Cultivar adaptation impacts on breadbasket location under warming 90
4.5 Changes in breadbasket location for N. American maize 91
4.6 Regional differences in breadbasket shift 95

5.1 Historical nitrogen intensity and yields 103
5.2 Historical simulation temperature-yield curves 110
5.3 Panel model estimation of temperature impacts including nitrogen 111
5.4 Potential mis-estimation of climate impacts on crops in nutrient limited areas . 112
5.5 Current nitrogen application rates and potential missed CO_2 benefit 113
5.6 Global grain production change under possible high-end climate change 114

S1 Presently cultivated area for irrigated crops in the real world 144
S2 Comparison across the three reanalysis products for precipitation 145
S3 Comparison across the three reanalysis products for temperature 146
S4 Figure 2.5a for all crops 150
S5 Figure 2.5b for all crops 151
S6 Figure 2.6a for all crops 152
S7 Figure 2.6b for all crops 153
S8 Figure 2.5a for all crops 154
S9 Figure 2.5a for all crops on only cultivated area 155
S10 Figure 2.5b for all crops on only cultivated area 156

S11 Figure 2.6a for all crops on only cultivated area . . . 157
S12 Figure 2.6b for all crops on only cultivated area . . . 158
S13 Figure 2.5a for all crops for A1 . . . 159
S14 Figure 2.5b for all crops for A1 . . . 159
S15 Figure 2.6b for all crops for A1 . . . 160
S16 Number of models providing simulations . . . 161
S17 Presently cultivated area . . . 162
S18 Change in within-growing-season daily minimum temperature variability . . . 164
S19 Figure S18 for daily maximum temperature . . . 165
S20 Figure S18 except now for maize only . . . 165
S21 Illustration of spatial variations in yield response for A1 . . . 166
S22 Illustration of variations in yield response across models for A1 . . . 166
S23 Emulation error for A1 high CO_2 . . . 167
S24 Emulation error for A1 low CO_2 . . . 168
S25 Emulation error for irrigation water demand . . . 168
S27 Figure 3.10 for individual models . . . 170
S28 Emulated damage functions for rainfed crops for A1 . . . 171
S29 Figure 3.11 for spring wheat . . . 172
S30 Illustration of the factors affecting yields in more realistic climate scenarios . . . 172
S31 Figure 3.4 for reduced form emulator . . . 174
S32 Figure 3.5 for reduced form emulator . . . 174
S33 Figure 3.5 for reduced form emulator . . . 175
S34 Figure 3.7 for reduced form emulator . . . 175
S35 Figure 3.8 for reduced form emulator . . . 176
S36 Figure 3.11 for reduced form emulator . . . 176
S37 Example of emulator failure due to interacting terms . . . 177
S38 Example of emulator failure due to abrupt change in yield . . . 178
S39 Spatial yield response and emulator error for LPJmL . . . 179
S40 Spatial yield response and emulator error for pDSSAT . . . 180
S41 Root mean squared error for cross validation for pDSSAT . . . 181
S42 Root mean squared error for cross validation for PROMET . . . 182
S43 Root mean squared error for cross validation for JULES . . . 183
S44 Root mean squared error for cross validation for EPIC-TAMU . . . 184
S45 Root mean squared error for cross validation for GEPIC . . . 185
S46 Root mean squared error for cross validation for PEPIC . . . 186
S47 Major Breadbasket locations . . . 190
S48 Presently cultivated area for rainfed crops . . . 191
S49 Ecosystem vs site based model comparison . . . 192
S50 CMIP6 temperature shifts along latitudes (SSP585) . . . 193
S51 CMIP6 temperature shifts along latitudes (SSP126) . . . 194
S52 Figure 4.2 for high CO_2 simulations . . . 195
S53 Figure 4.2 for irrigated simulations . . . 195
S54 Figure 4.2 for the pDDSST model . . . 196

S55 Figure 4.2 for the pDDSST model high CO_2 . . . 196
S56 Figure 4.2 for the LPJmL model . . . 197
S57 Figure 4.2 for the LPJmL model high CO_2 . . . 197
S58 Figure 4.2 for the PEPIC model . . . 198
S59 Figure 4.2 for the PEPIC model high CO_2 . . . 198
S60 Figure 4.2 for the EPIC-TAMU model . . . 199
S61 Figure 4.2 for the EPIC-TAMU model high CO_2 . . . 199
S62 Figure 4.2 for the PROMET model . . . 200
S63 Figure 4.2 for the PROMET model high CO_2 . . . 200
S64 Figure 4.2 for the LPJ-GUESS model . . . 201
S65 Figure 4.2 for the LPJ-GUESS model high CO_2 . . . 201
S66 Yield centroid for each model for maize under A0 growing season . . . 202
S67 Yield centroid for each model for maize under A0 growing season . . . 203
S68 Figure 4.3 with simulations at T+6°C and CO_2 of 810 ppm . . . 204
S69 Figure 4.4 for 2°C warming . . . 205
S70 Figure 4.4 for high CO_2 . . . 206
S71 Climate model uncertainty . . . 207
S72 Figure 4.2 for LPJmL HadGEM transient runs . . . 208
S73 Figure 4.2 for LPJmL MIROC transient runs . . . 209
S74 Figure 4.2 for LPJmL IPSL transient runs . . . 210
S75 Figure 4.2 for LPJmL GFDL transient runs . . . 211
S76 Latitudinal distribution of the highest yielding grid cells . . . 212
S77 Figure 4.6 for maize in Africa . . . 213
S78 Latitudinal uncertainty in growing seasons changes . . . 214
S79 Current global nitrogen applications . . . 215
S80 Nitrogen-driven climate impact miss-estimation by main KG region . . . 215

LIST OF TABLES

2.1 GGCMI Phase 2 input levels for the parameter sweep 11
2.2 Protocol output variables 19
2.3 Models included in GGCMI Phase 2 22

3.1 Crop models included in GGCMI Phase 2 emulators 42
3.2 RMSE of emulator replication of simulated yields of rainfed crops 62

4.1 The 7 process-based globally gridded crop models used in the study 79
4.2 The process-based crop model scenarios presented in this analysis. 80

S1 Key model details 147
S2 Change in growing season daily maximum temperature variability 163
S3 Growing season changes for the LPJmL simulations 188
S4 Figure 4.6 summary statistics 189
S5 Percent of area within the breadbasket 189

ACKNOWLEDGMENTS

This work was supported by:

- the University of Chicago Center for Robust Decision-making on Climate and Energy Policy (RDCEP) which is funded by NSF grant SES-1463644 through the Decision Making Under Uncertainty program
- the NSF National Research Traineeship program (grant DGE-1735359)
- the NSF Graduate Research Fellowship Program (grant DGE-1746045)

STATEMENT OF ORIGINALITY AND COPYRIGHT ATTRIBUTION

All text and analysis in Chapters 2, 3, and 4 was produced in collaboration with the authors listed at the beginning of each chapter with James Franke serving as the lead author in each case. All figures contained within were produced by James Franke using data from the publicly available sources listed in each section. All figures and text in Chapter 2, 3, and 4 have been published at the following locations:

- **Franke, J. A.**, Müller, C., Elliott, J., Ruane, A. C., Jägermeyr, J., Balkovic, J., Ciais, P., Dury, M., Falloon, P. D., Folberth, C., Francois, L., Hank, T., Hoffmann, M., Izaurralde, R. C., Jacquemim, I., et al. 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* (2020). *https://doi.org/10.5194/gmd-13-2315-2020*. Published under Creative Commons Attribution 4.0 International copyright: *https://creativecommons.org/licenses/by/4.0/*.
- **Franke, J. A.**, Müller, C., Elliott, J., Ruane, A. C., Jägermeyr, J., Snyder, A., Dury, M., Falloon, P. D., Folberth, C., Francois, L., Hank, T., Izaurralde, R. C., Jacquemin, I., Jones, C., Li, M., et al. The GGCMI Phase 2 emulators: global gridded crop model responses to changes in CO_2, temperature, water, and nitrogen (version 1.0) *Geoscientific Model Development* (2020). *https://doi.org/10.5194/gmd-13-3995-2020*. Published under Creative Commons Attribution 4.0 International copyright: *https://creativecommons.org/licenses/by/4.0/*.
- **Franke, J. A.**, Müller, C., Minoli, S., Elliott, J., Folberth, C., Gardner, C., Hank, T., Izaurralde, R. C., Jägermeyr, J., Jones, C. D., Liu, W., Olin, S., Pugh, T. A., Ruane, A. C., Stephens, H., et al. Agricultural breadbaskets shift poleward given adaptive farmer behavior under climate change. *Global Change Biology* (2021). *https://doi.org/10.1111/gcb.15868*. Wiley holds the rights but grants permission to authors for full reproduction for their dissertation.

All text and analysis in Chapter 5 was produced by James Franke.

ABSTRACT

Importance: Understanding global agricultural response to a changing climate alongside changes in land use, technology and food demand is necessary for future food security. Global, spatially resolved projections of major agricultural products remain very uncertain however; resulting from uncertainties in economic development, climate sensitivity, crop response to climate, future management changes, and genetics. Highlighting the spatial disparities in climate impacts on agriculture while improving projection methodologies will help target adaptation efforts, and guide investment and intervention to build a more resilient food system. The following chapters aim: (1) to improve the simulation of global crop production under changing climate, (2) to build crop model emulators that can be utilized for economic impact assessments (3) to better understand how cultivation regions for major crops may shift in the future, and (4) to assess uncertainties around interactions between applied nutrients and climate.

Approach: This work is based on a set of harmonized, globally gridded, processed-based crop model simulations over a set of nutrient (atmospheric [CO_2] and applied nitrogen fertilizers) and climate (daily temperature, precipitation, solar radiation, etc) input combinations. This set of simulations is collectively known as the Globally Gridded Crop Model Intercomparison (GGCMI) project Phase 2 under the Agriculture Model Intercomparison and Improvement Project (AgMIP) umbrella. Chapter 2 describes the Phase 2 experiment design, presents overall simulation results, and Chapter 3 presents the development and evaluation of a set of crop model emulators designed to be used in economic frameworks. Chapter 4 employs these same crop models and some additional, targeted simulations to test how regions of peak yield may shift under possible climate change. In Chapter 5, I utilize the Phase 2 simulations and emulators to test how the most important nutrient, applied nitrogen fertilizers, modulates the climate response of agriculture for the major grains globally.

Key Findings: Warming uniformly reduces crop production across models in both rain-fed

and irrigated systems for maize, rice, wheat and soybeans in all but the highest latitude regions. Presently hotter, dryer regions are more vulnerable to warming with tropical latitudes suffering greater simulated losses under possible climate change. [CO_2] direct fertilization compensates for temperature losses and often completely reverses the signal – especially for crops other than maize. The mean climatological yield response to climate change in process-based models can be faithfully represented by statistical emulation around the mean growing season temperature change or mean precipitation change. Changes in temperature variability matter little at the mean-climatological timescale to yield changes in process based models, with some potential exceptions at the highest latitudes. With adapted growing seasons, Regions of peak yield in the extra-tropics shift poleward substantially under expected climate change. Nitrogen inputs modulate climate impacts dramatically, with grains under high input showing amplified sensitivity to both temperature and [CO_2] changes. Climate change crop yield impacts are uncertain everywhere globally –with more pessimistic projections in the tropics being the major consistent feature across models and scenarios.

CHAPTER 1

INTRODUCTION

Understanding agricultural production response to a changing climate is critically important. The global food production system will face pressures from sharply increased demand over the coming decades [45, 25, 95, 57]; therefore climate-related reductions in supply could have severe socioeconomic consequences [e.g. 44, 210, 16, 189]. Multiple studies using different crop or climate models concur in projecting yield reductions on currently cultivated cropland under *business-as-usual* climate scenarios though yield projections can show considerable uncertainty [e.g. 165, 22, 178, 124, and references therein]. I seek to add to this research by highlighting some spatial aspects of global impacts of climate on food production.

The scale of the food production system makes studying it a geoscience problem. No human activity impacts the earth system as much as agriculture [162]. With ice excluded, nearly 60% [153, 46] of the land surface of the Earth (including deserts) is currently employed in agricultural production (food, fodder, biofuel, and pasture). Three-quarters of all freshwater use (and as much as as much as 90% of *consumptive* uses) [147] is agricultural. As much as a third of GHG emissions [202, 31] are agricultural (the exact fraction is uncertain and depends on how much food transport is included, how industrial nitrogen fertilizer production and resultant NO_x emissions are allocated, how land-use change emissions are accounted for, and how the GWP is calculated for methane). No other economic sector may be impacted as much by climate change as agriculture [e.g. 137], with countries most reliant on agriculture experiencing the greatest loss in GDP under expected climate change [e.g. 98]. Projecting global agricultural output under climate change is therefore not strictly an ecology, biology, or agronomy problem–but a geoscience problem.

However, projecting future global food production is hampered by numerous uncertainties. Inter-annual variability illustrates that weather–and therefore climate–is critical to yields [156], but response to changes in temperature and precipitation are both generally

uncertain [though temperature dominates in this regard, 113]. Spatial heterogeneity can be large with the exact characteristics of soil chemistry and drainage (e.g. slope, substrate) often acting on the 10-50-meter scale [9, 17]. Historical data is extremely poor at relatively high resolution, especially in countries other than the U.S. and Western Europe. On top of data challenges, historical dynamism in technology, cultivar genetics, and management intensification have driven production increases that dwarf inter-annual variability in most regions. Production has tripped globally in the past 50 years [103] on relatively constant agricultural land [44]. (Agricultural land is actually decreasing in the U.S. at a rate of about 1 hectare every 2 minutes [199].) Agriculture will likely continue to change in the future as management practices, genetics, and climate continue to move into uncharted waters.

Information about crop response to climate in heterogeneous environments and under technological changes comes from two distinct sources: Data from natural experiments where historical weather variability translates into historical yield variability and data from controlled experiments in fields or greenhouses where conditions are intentionally imposed on plants by researchers. Existing models for projecting future crop yields also fall into two categories: *Statistical models* relate historical weather to historical reported yields and are built using these natural experiments and *Process-based models* –numerical simulations based in biological and ecological principles mainly derived from controlled experiments. (Ecosystem models, a subset of the process-based models, typically also contain information derived from larger geospatial data sets and some statistical models incorporate data from controlled field studies.) Both types of models are routinely used, and comparative studies have concluded that when done carefully, both approaches can provide similar yield estimates [e.g. 114, 124, 112, 108, 161].

Process based models have been in use in research since the late 1950s [68, 27], and have steadily grown in complexity and are run today with considerable computing time on super-computers. Process-based crop models are necessary tools for climate-change impact

assessments on agriculture since they allow representing future climate and management conditions not sampled in the historical record, including higher [CO_2] levels and new locations to which cultivation may shift [e.g. 150, 161, 121]. Process-based models have therefore been widely used in studies on future food security [209, 35, 50], options for climate mitigation [127] and adaptation [24], and future sustainable development [78, 87]. Process-based models also enable global, gridded simulations for understanding the global dynamics of agricultural trade, including cultivation area changes and crop-selection switching under climate change [166, 169], because global market mechanisms may strongly modulate climate-change impacts [189, 64].

Modeling crop responses with process-based models continues to be challenging as crop growth is a function of complex interactions between climate inputs, soil, and management practices [e.g. 167, 18]. Process-based models often disagree widely [126], especially in developing areas [146] where statistical models also perform the worst. Management and genetics must be explicitly modeled at whatever scale they differ. For example, simply getting growing season dates correct (to say nothing of genetics, the exact timing of fertilizer applications, or the details of pesticide or herbicide use) is extremely challenging and can make or break the simulation accuracy [87]. Nevertheless, models tend to agree broadly in major response patterns, including a reasonable representation of the spatial pattern in historical yields of major crops and projections of shifts in yield under future climate scenarios [e.g. 34, 126]. But process-based models still struggle with some important details, including reproducing historical year-to-year variability in many regions [e.g. 126], reproducing historical yields when driven by reanalysis weather [e.g. 55], and low sensitivity to extreme events [e.g. 56, 179]. Projections therefore have considerable uncertainty [90, 80, 3, 5, 4].

Model intercomparison projects such as the Agricultural Model Intercomparison and Improvement Project [AgMIP, 163] are crucial in quantifying uncertainties in model projections [165]. Intercomparison projects have also been used to develop protocols for evaluating

overall model performance [34, 126] and to assess the representation of individual physical mechanisms such as water stress and [CO_2] fertilization [e.g. 178]. However, to date, few projects have systematically sampled critical factors that may interact strongly in affecting crop yields. A number of modeling exercises in the past five years have begun to use systematic parameter sweeps in crop model evaluation and emulation [e.g. 173, 143, 51, 188, 176], but all involve limited sites, and many have limited crops and scenarios. Global evaluations at high spatial resolution are needed.

CHAPTER 2

THE GGCMI PHASE 2 EXPERIMENT: GLOBAL GRIDDED CROP MODEL SIMULATIONS UNDER UNIFORM CHANGES IN CO_2, TEMPERATURE, WATER, AND NITROGEN LEVELS

Published in *Geoscientific Model Development* in collaboration with:

Christoph Müller[3], Joshua Elliott[2,4], Alex C. Ruane[5], Jonas Jägermeyr[2,3,4,5], Juraj Balkovic[6,7], Philippe Ciais[8,9], Marie Dury[10], Pete Falloon[11], Christian Folberth[6], Louis François[10], Tobias Hank[12], Munir Hoffmann[13,22], R. Cesar Izaurralde[14,15], Ingrid Jacquemin[10], Curtis Jones[14], Nikolay Khabarov[6], Marian Koch[13], Michelle Li[2,16], Wenfeng Liu[8,17], Stefan Olin[18], Meridel Phillips[5,19], Thomas A. M. Pugh[20,21], Ashwan Reddy[14], Xuhui Wang[8,9], Karina Williams[11], Florian Zabel[12], and Elisabeth Moyer[1,2]

[1]Department of the Geophysical Sciences, University of Chicago, Chicago, IL, USA [2]Center for Robust Decision-making on Climate and Energy Policy (RDCEP), University of Chicago, Chicago, IL, USA [3]Potsdam Institute for Climate Impact Research, Member of the Leibniz Association, Potsdam, Germany [4]Department of Computer Science, University of Chicago, Chicago, IL, USA [5]NASA Goddard Institute for Space Studies, New York, NY, United States [6]Ecosystem Services and Management Program, IIASA, Laxenburg, Austria [7]Department of Soil Science, Comenius University in Bratislava, Bratislava, Slovak Republic [8]Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, 91191 Gif-sur-Yvette, France [9]Sino-French Institute of Earth System Sciences, Peking University, Beijing, China [10]Unité de Modélisation du Climat et des Cycles Biogéochimiques, UR SPHERES, University of Liège, Belgium [11]Met Office Hadley Centre, Exeter, United Kingdom [12]Department of Geography, Ludwig-Maximilians-Universität, Munich, Germany [13]Georg-August-University Göttingen, Göttingen, Germany [14]Department of Geographical Sciences, University of Maryland, College Park, MD, USA [15]Texas Agrilife Research and Extension, Texas A&M University, Temple, TX, USA [16]Department of Statistics, University of Chicago, Chicago, IL, USA [17]EAWAG, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland [18]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden [19]Earth Institute Center for Climate Systems Research, Columbia University, New York, NY, USA [20]School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK. [21]Birmingham Institute of Forest Research, University of Birmingham, Birmingham, UK. [22]Leibniz Centre for Agricultural Landscape Research (ZALF), D-15374 Müncheberg, Germany

Concerns about food security under climate change motivate efforts to better understand future changes in crop yields. Process-based crop models, which represent plant physiological and soil processes, are necessary tools for this purpose since they allow representing future climate and management conditions not sampled in the historical record and new locations to which cultivation may shift. However, process-based crop models differ in many critical

details, and their responses to different interacting factors remain only poorly understood. The Global Gridded Crop Model Intercomparison (GGCMI) Phase 2 experiment, an activity of the Agricultural Model Intercomparison and Improvement Project (AgMIP), is designed to provide a systematic parameter sweep focused on climate change factors and their interaction with overall soil fertility, to allow both evaluating model behavior and emulating model responses in impact assessment tools. In this paper we describe the GGCMI Phase 2 experimental protocol and its simulation data archive. Twelve crop models simulate five crops with systematic uniform perturbations of historical climate, varying [CO_2], temperature, water supply, and applied nitrogen ("CTWN") for rainfed and irrigated agriculture, and a second set of simulations represents a type of adaptation by allowing the adjustment of growing season length. We present some crop yield results to illustrate general characteristics of the simulations and potential uses of the GGCMI Phase 2 archive. For example, in cases without adaptation, modeled yields show robust decreases to warmer temperatures in almost all regions, with a nonlinear dependence that means yields in warmer baseline locations have greater temperature sensitivity. Inter-model uncertainty is qualitatively similar across all the four input dimensions, but is largest in high-latitude regions where crops may be grown in the future.

2.1 Introduction

Understanding crop yield response to a changing climate is critically important, especially as the global food production system will face pressure from increased demand over the next century [43, 16]. Climate-related reductions in supply could therefore have severe socioeconomic consequences [e.g. 189, 210]. Multiple studies using different crop or climate models concur in projecting sharp yield reductions on currently cultivated cropland under business-as-usual climate scenarios, although their yield projections show considerable spread [e.g. 165, 178, and references therein]. Although forecasts of future yields reductions can be

made with simple statistical models based on regressions in historical weather data, process-based models, which simulate the effect of temperature, water and nutrient availability, and atmospheric [CO_2] concentration on the process of photosynthesis and the biology and phenology of individual crops, play a critical role in assessing the impacts of climate change.

Process-based models are necessary for understanding crop yields in novel conditions not included in historical data, including higher [CO_2] levels, out-of-sample combinations of rainfall and temperature, cultivation in areas where crops are not currently grown, and differing management practices [e.g. 150, 161, 120]. Process-based models have therefore been widely used in studies on future food security [209, 35, 50], options for climate mitigation [126] and adaptation [24], and future sustainable development [78, 89]. They are a necessity for global gridded simulations, which allow understanding the global dynamics of agricultural trade, because global market mechanisms can strongly modulate the economic impacts of regional yield changes [189, 64]. Global simulations are especially necessary in studying agricultural effects of climate change [126], since systematic climate assessments must account for cultivation area changes and crop selection switching [166, 169] and must consider inter-regional differences [e.g. 132, 210].

Modeling crop responses, however, continues to be challenging, as crop growth is a function of complex interactions between climate inputs, soil, and management practices [18, 167]. Models tend to agree broadly in major response patterns, including a reasonable representation of the spatial pattern in historical yields of major crops and projections of shifts in yield under future climate scenarios [e.g. 34, 126]. But process-based models still struggle with some important details, including reproducing historical year-to-year variability in many regions [e.g. 126, 87], reproducing historical yields when driven by reanalysis weather [e.g. 55], and low sensitivity to extreme events [e.g. 56, 179]. Global models pose additional challenges due to variable input data quality and limited ability for model calibration. Long-term projections therefore retain considerable uncertainty [213, 90, 80, 3, 5, 4].

Model intercomparison projects such as the Agricultural Model Intercomparison and Improvement Project [AgMIP, 163] are crucial in quantifying uncertainties in model projections [165]. Intercomparison projects have also been used to develop protocols for evaluating overall model performance [34, 126] and to assess the representation of individual physical mechanisms such as water stress and [CO_2] fertilization [e.g. 178]. However, to date, few such projects have systematically sampled critical factors that may interact strongly in affecting crop yields. A number of modeling exercises in the last five years have begun to use systematic parameter sweeps in crop model evaluation and emulation [e.g. 173, 143, 51, 188, 176], but all involve limited sites and most also limited crops and scenarios.

The Global Gridded Crop Model Intercomparison (GGCMI) Phase 2 experiment is the first global gridded crop model intercomparison involving a systematic parameter sweep across critical interacting factors. GGCMI Phase 2 is an activity of AgMIP, and a continuation of a multi-model comparison exercise begun in 2014. The initial GGCMI Phase 1 [34, 126] compared harmonized yield simulations over the historical period, with primary goals of model evaluation and understanding sources of uncertainty (including model parameterization, weather inputs, and cultivation areas). See also [41] and [146] for more information. GGCMI Phase 2 compares simulations across a set of inputs with uniform perturbations to historical climatology, including [CO_2], temperature, precipitation, and applied nitrogen, as well as adaptation to shifting growing seasons (collectively referred to as "CTWN-A"). The CTWN-A experiment is inspired by AgMIP's Coordinated Climate-Crop Modeling Project [C3MP, see 173, 118] and contributes to the AgMIP Coordinated Global and Regional Assessments [CGRA, see 166, 169].

In this paper, we describe the GGCMI Phase 2 model experiments and present initial summary results. In the sections that follow, we describe the experimental goals and protocols; the different process-based models included in the intercomparison; the levels of participation by the individual models. We then provide an assessment of model fidelity based

on observed yields at the country level, and show some selected examples of the simulation output dataset to illustrate model responses across the input dimensions.

2.2 Simulation Objectives and Protocol

2.2.1 Goals

The guiding scientific rationale of GGCMI Phase 2 is to provide a comprehensive, systematic evaluation of the response of process-based crop models to critical interacting factors, including [CO_2], temperature, water, and applied nitrogen under two contrasting assumptions on growing season adaptation (CTWN-A). The dataset is designed to allow researchers to:

- Enhance understanding of models' sensitivity to climate and nitrogen drivers.
- Study the interactions between climate variables and nitrogen inputs in driving modeled yield impacts.
- Characterize differences in crop responses to climate change across the Earth's climate regions.
- Provide a dataset that allows statistical emulation of crop model responses for downstream modelers.
- Explore the potential effects on future yield changes of adaptations in growing season length.

2.2.2 Modeling Protocol

The GGCMI Phase 1 intercomparison was a relatively limited computational exercise, requiring yield simulations for 19 crops across a total of 310 model-years of historical scenarios, and had the participation of 14 modeling groups. The GGCMI Phase 2 protocol is substantially

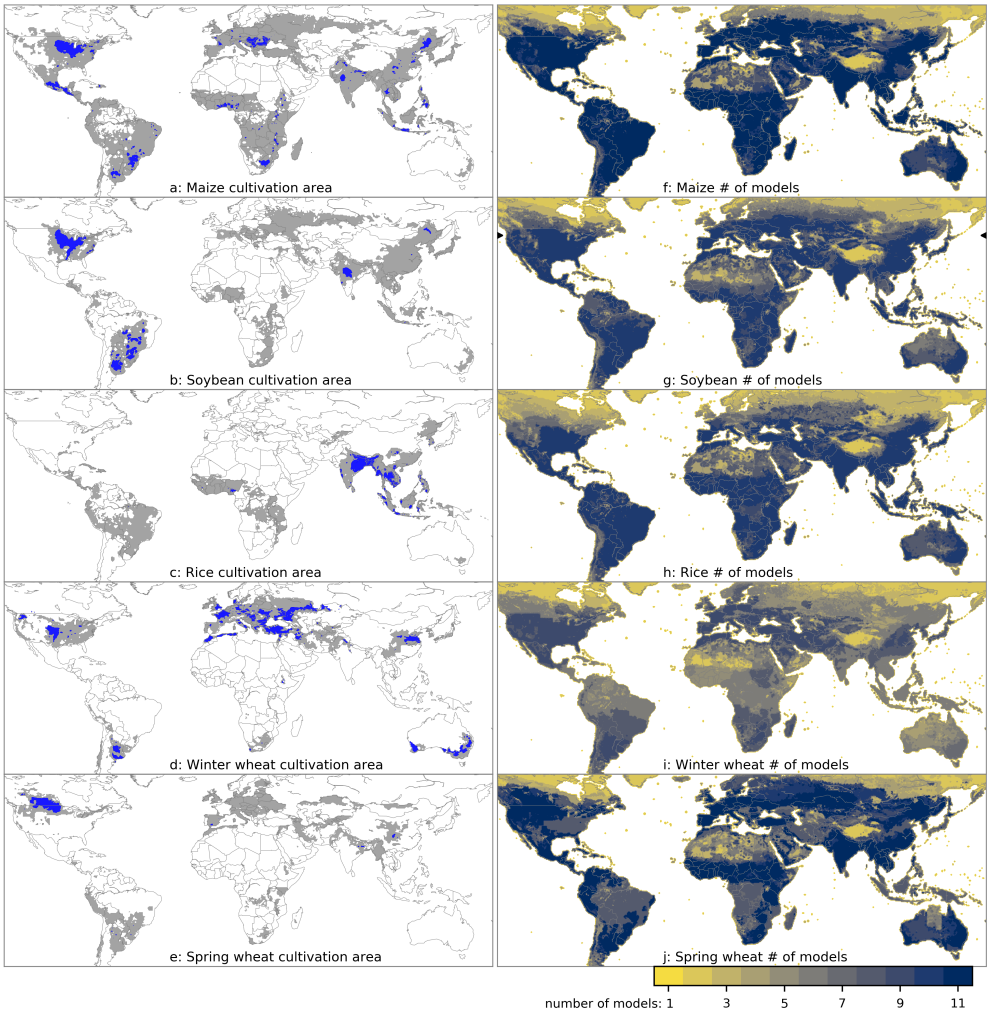


Figure 2.1: GGCMI Phase 2 simulation coverage and global cropland. Cultivated areas for maize, rice, and soybean from the MIRCA2000 ("Monthly Irrigated and Rainfed Crop Areas around the year 2000") dataset [145]. Blue indicates grid cells with more that 20,000 hectares (10% of the equatorial grid cell) and gray contour shows gridcells with more that 10 hectares cultivated. Areas for winter and spring wheat areas are adapted from MIRCA2000 and two other sources; see text for details. For irrigated crops, see supplemental Figure S1. **Right panel:** Number of models providing simulations for each grid cell. All models provide the minimum areal coverage of the GGCMI Phase 2 protocol, but some provide extra coverage at high latitudes or in arid or otherwise unsuitable areas. Source: [47].

larger, involving over 1400 individual 30-year global scenarios, or over 42,000 model-years; 12 modeling groups nevertheless participated. To reduce the computational load, the GGCMI Phase 2 protocol reduces the number crops to 5 (maize, rice, soybean, spring wheat, and winter wheat). The reduced set of crops includes the three major global cereals and the

major legume and accounts for over 50% of human calories in 2016: nearly 3.5 billion tons or 32% of total global crop production by weight [46]. This set of major crops has the advantage of historical yield data globally available at sub-national scale [156, 81], and has been frequently used in subsequent analyses [e.g. 126, 146].

The Phase 2 protocol involves a suite of uniform perturbations from a historical climate timeseries. The baseline climate scenario for GGCMI Phase 2 is one of the weather products used in Phase 1, daily climate inputs for 1980-2010 from the 0.5 degree NASA AgMERRA ("Agricultural"-modified Modern Era Retrospective analysis for Research and Applications) gridded re-analysis product. AgMERRA is specifically designed for agricultural modeling, with satellite-corrected precipitation [171]. The experimental protocol consists of 9 levels for water supply perturbations, 7 for temperature, 4 for [CO_2], and 3 for applied nitrogen, for a total of 756 simulations (Table 2.1), 672 for rainfed agriculture and an additional 84 for irrigated (W_∞). Values of climate variable perturbations are selected to represent reasonable ranges for changes over the medium term (to 2100) under business-as-usual emissions. Values for nitrogen application levels are intended to cover a wide range of potentials. The resulting GGCMI Phase 2 dataset captures the distribution of crop model responses over a wide range of potential future climate and management conditions.

Table 2.1: GGCMI Phase 2 input levels for the parameter sweep. Values for temperature and water supply are perturbations from the historical climatology. For water supply, perturbations are fractional changes to historical precipitation, except in the irrigated (W_∞) simulations, which are all performed with the maximum beneficial levels of water. Bold font indicates the 'baseline' historical level. The full protocol samples across all parameter combinations for a total of 756 cases.

Input variable	**Tested range**	**Unit**
[CO_2](C)	**360**, 510, 660, 810	ppm
Temperature (T)	-1, **0**, 1, 2, 3, 4, 6	°C
Precipitation (W)	-50, -30, -20, -10, **0**,	%
	10, 20, 30, (and W_∞)	
Applied nitrogen (N)	10, 60, **200**	kg ha^{-1}
Adaptation (A)	**A0: none**, A1: maintain original growing season length	-

The protocol samples over all possible permutations of individual perturbations, i.e. all values are applied across all crops and regions, so that the protocol includes many combinations that are not realistic. For example, we simulate high N application to soybeans, which are N-fixers and need little fertilizer. This choice also means that $[CO_2]$ changes are applied independently of changes in climate variables, so that higher $[CO_2]$ is not associated with higher temperatures or other particular climate changes. The purpose of the experiment is not to produce individual scenarios that represent realistic future states, but to sample over a wide range of parameter space to enable understanding the factors that drive agricultural changes.

While all CTWN perturbations are applied uniformly across the historical timeseries, they are applied in different ways. $[CO_2]$ and nitrogen levels are specified as discrete values applied uniformly over all grid cells. Temperature perturbations are applied as absolute offsets from the daily mean, minimum, and maximum temperature timeseries for each grid cell, and water perturbations are applied as fractional changes to daily precipitation. The irrigated scenario (W_∞) is a particular case of water supply levels, in which crops are assumed to have no water constraints. That is, all crop water requirements are fulfilled regardless of local water supply limitations. To facilitate comparison, irrigated simulations use the same growing seasons as all other simulations, even though in reality irrigated growing seasons may be different [145], and both irrigated and rainfed cases are simulated with near-global coverage.

The uniform perturbations of the GGCMI Phase 2 protocol require some care in interpretation. Temperature and precipitation perturbations should be considered as differences from historical climatology within the growing season only. That is, a T+1 simulation represents a 1 °C warmer growing season, not a 1 °C warmer annual mean temperature. (The distinction is important because in climate projections, winters generally warm more than summers [e.g. 65].) In the GGCMI Phase 2 protocol, temperature and precipitation pertur-

bations are applied uniformly in space, but future changes in temperature and precipitation will not be spatially or temporally uniform. In a realistic climate projection, higher latitudes generally warm more strongly than lower latitudes [e.g. 61], and the northern high latitudes warm more quickly than the southern ones. A GGCMI Phase 2 simulation therefore represents a possible future state that *could* occur in each grid cell, but not one that would in reality occur simultaneously in all grid cells across the globe. The GGCMI Phase 2 simulations are intended to be used for climate impact assessment not directly but instead as a "training set" for statistical emulation of each crop model. Once an emulator is constructed from the outputs described here, it can be driven with growing-season climate anomalies from any climate model projection. The GGCMI Phase 2 protocol does not involve any simulated changes in climate variability, but [48] demonstrate that these effects are relatively minor and that GGCMI Phase 2 emulators can effectively reproduce crop model yields under realistic future climate scenarios.

The area simulated in the GGCMI Phase 2 protocol extends considerably outside currently cultivated areas, because cultivation may shift under climate change. Figure 2.1 shows both the present-day cultivated area of rainfed crops (left) and model coverage (right). (See Supplemental Figure S1-2 for currently cultivated area for irrigated crops; model coverage is the same.) Each model covers all currently cultivated areas and much of the uncultivated land area, run at 0.5 degree spatial resolution. To reduce the computational burden, the protocol requires simulation over only 80% of Earth land surface area, omitting areas assumed to remain non-arable even under an extreme climate change, including Greenland, far-northern Canada, Siberia, Antarctica, the Gobi and Sahara Deserts, and Central Australia. The protocol also allows omitting regions judged unsuitable for cropland for non-climatic reasons. Selection criterion involve a combination of soil suitability indices at 10 arc-minute resolution and excludes those 0.5 degree grid cells in which at least 90% of the area is masked as unsuitable according to any single index, and which do not contain any currently cultivated

cropland. Currently cultivated areas are provided by the MIRCA2000 (Monthly Irrigated and Rainfed Crop Area) data product [145]. Soil suitability indices measure excess salt, oxygen availability, rooting conditions, toxicities, and workability, and are provided by the IIASA (International Institute for Applied Systems Analysis) Global Agro-Ecological Zone model [GAEZ, 40]. The procedure follows that proposed by [150]. All modeling groups simulate the minimum required coverage, but some provide simulations that extend into masked zones, including e.g. the Sahara Desert and Central Australia (Figure 2.1, right).

2.2.3 Harmonization Between Models

The 12 models included in GGCMI Phase 2 are all process-based crop models that are widely used in impacts assessments (Table 2.3). Although some models share a common base (e.g. the LPJ or EPIC families of models), they have subsequently developed independently. Wherever possible, the GGCMI Phase 2 protocol harmonizes inputs, but differences in model structure mean that several key factors cannot be fully standardized across the experiment. These include soil treatment (which affects soil organic matter and carry-over effects of soil moisture across growing years) and baseline climate inputs.

While 10 of the 12 models participating in GGCMI Phase 2 use the AgMERRA historical daily climate data product, two models require sub-daily input data and thus use different baseline climate inputs: PROMET uses ERA-Interim reanalysis [30], and JULES uses a bias-corrected version of ERA-Interim, the 3-hour WFDEI (WATCH-Forcing-Data-ERA-Interim) [208], specifically the WFDEI version with precipitation bias-corrected against the CRU TS3.101/TS3.21 precipitation totals [63]. The data products show some differences (Figures S3-S4, which compare data products over currently cultivated areas for each crop). For example, for maize-growing areas, ERA-Interim daily precipitation is biased high from that in AgMERRA by 7% ($<$ 1 sigma), while mean daily precipitation in WFDEI is only 3% higher. Precipitation differences are largest in wheat areas, where ERA-Interim is sub-

stantially wetter (+60 mm year^{-1} or 10%). Temperature differences are largest for rice, with ERA-Interim 1°C cooler than AgMERRA, and smaller for other crops, e.g. maize with ERA-Interim 0.45°C cooler and WFDEI 0.1°C warmer. These differences are relatively small compared to the perturbations tested in the protocol.

Planting dates and growing season lengths are standardized across models, following the procedure described in [34] for the *fullharm* setting. (The exception is that Phase 2, unlike Phase 1, uses identical growing seasons for rainfed and irrigated cases, to allow for direct comparison of simulations along the W dimension.) This harmonization is important because the parametrization of growing seasons can have strong effects on simulated yields [126, 87]. In all the GGCMI Phase 2 crop models, sowing dates are prescribed directly, but the length of the growing season is a product of crop phenology, which is driven mostly by phenological parameters and temperature. Modelers were therefore asked to adjust their phenological parameters so that the average growing season length of the baseline scenario (C=360, T=0, W=0) matched the harmonization target. (The one exception to this harmonization protocol involves CARAIB, whose team kept their own growing season specifications rather than tuning to standard lengths.) Two aspects of the procedure should be noted. First, the target growing seasons used in GGCMI Phase 2 are crop- and location-specific. For example, present-day maize is sown in March in Spain, in July in Indonesia, and in December in Namibia [145]. Second, because temperature varies between years in the 30-year baseline climatology, realized growing season length will still vary in individual years even after harmonization.

The dependence of harvest dates on climate parameters means that growing seasons will alter under climate change in a model with phenological parameters tuned to match target growing seasons in the baseline climate. In general, warmer future scenarios produce shorter growing seasons. We denote simulations that allow these future changes as "A0" experiments, where 0 denotes "no adaptation". The GGCMI Phase 2 protocol includes a

second set of experiments, "A1", that assume that future cultivars are modified to adjust to changes along the T dimension in the CTWN experiment. For these simulations, modelers adjust phenological parameters for each temperature scenario to hold growing season length approximately constant. (CARAIB simulations follow the same principle, fixing growing season length at their baseline levels.) That is, the A1 simulations require running a model with seven different choices of cultivar parameters, one per warming level. Parameter settings for T=0 are identical in both A0 and A1. The A1 simulations roughly capture the case in which adaptive crop cultivar choice ensures that crops reach maturity at roughly the same time as in the current temperature regime. This assumption is simplistic, and does not reflect realistic opportunities and limitations to adaptation [201, 24], but provides some insight into how crop modifications could alter projected impacts on yields and is sufficiently easy to implement in a large model intercomparison project as GGCMI.

Growing seasons for maize, rice, and soybean are taken from the SAGE (Center for Sustainability and the Global Environment, University of Wisconsin) crop calendar [177], gap-filled with the MIRCA2000 crop calendar [145] and, if no SAGE or MIRCA2000 data are available, with simulated LPJmL growing seasons [204] and are identical to those used in GGCMI Phase 1 [34]. In GGCMI Phase 2, we separately treat spring and winter wheat and so must define different growing seasons for each. As for the other crops, we use the SAGE crop calendar, which separately specifies spring and winter wheat, as the primary source for 69% of grid cells. In the remaining areas where no SAGE information is available, we turn to, in order of preference, the MIRCA2000 crop calendar [145] and to simulated LPJmL growing seasons [204]. These datasets each provide several options for wheat growing season for each grid cell, but do not label them as spring or winter wheat. We assign a growing season to each wheat type for each location based on its baseline climate conditions. A growing season is assigned to winter wheat if all of the following hold, and to spring wheat otherwise:

- the monthly mean temperature is below freezing point ($<0°$C) at most for 5 months

per year (i.e. winter is not too long)

- the coldest 3 months of a year are below 10°C (i.e. there is a winter)
- the season start date fits the criteria that:
  - if in the N. hemisphere, it is after the warmest *or* before the coldest month of the year (as winter is around the end/beginning of the calendar year)
  - if in the S. hemisphere, it is after the warmest *and* before the coldest month of the year (as winter is in the middle of the calendar year)

Nitrogen (N) application is standardized in timing across models. N fertilizer is applied in two doses, as is often the norm in actual practice, to reduce losses to the environment. In the GGCMI Phase 2 protocol, half of the total fertilizer input is applied at sowing and the other half on day 40 after sowing, for all crops except for winter wheat. For winter wheat, in practice the application date for the second N fertilizer application varies according to local temperature, because the length of winter dormancy can vary strongly. In the GGCMI Phase 2 protocol, the second fertilization date for winter wheat must lie at least 40 days after planting and – if not contradicting the distance to planting – no later than 50 days before maturity. If those limits permit, the second fertilization is set to the middle day of the first month after sowing that has average temperatures above 8°C.

All stresses in models are disabled other than those related to nitrogen, temperature, and water. For example, model responses to alkalinity, salinity, and non-nitrogen nutrients are all disabled. No other external N inputs are permitted – that is, there is no atmospheric deposition of nitrogen – but some models allow additional release of plant-available nitrogen through mineralization in soils. In LPJmL, LPJ-GUESS and APSIM, soil mineralization is a part of model treatments of soil organic matter and cannot be disabled. Some additional differences in model structure mean that several key factors are not standardized

across the experiment. For example, carry-over effects across growing years including residue management and soil moisture are treated differently across models.

2.2.4 Output Data Products

All models in GGCMI Phase 2 provide 30-year timeseries of annual crop yields for each scenario, 0.5 degree grid cell and crop, in units of tons ha^{-1} year^{-1}. They also provide all available variables of the following 6: total above ground biomass yield; the dates of planting, anthesis and maturity; applied irrigation water in irrigated scenarios; and total evapotranspiration. We term these 7 variables the "mandatory" outputs, but note that some models do not compute all of them, e.g. CARAIB does not compute the anthesis date. Besides these 7 "mandatory" data products (Table 2.2, bold), the protocol requests any or all of 18 "optional" additional output variables (Table 2.2, plain text). Participating modeling groups provided between 3 (PEPIC) and 18 (APSIM-UGOE) of these optional variables.

All output data is supplied as netCDF version 4 files, each containing values for one variable in a 30-year timeseries associated with a single scenario, for all grid cells. File names follow the naming conventions of GGCMI Phase 1 [34], which themselves are derived from those of ISIMIP [50]. File names are specified as

$$[model]_[climate]_hist_fullharm_[variable]_[crop]_global_annual_[start-year]_[end-year]_[C]_[T]_[W]_[N]_[A].nc4$$

Here [*model*] is the crop model name; [*climate*] is the original climate input dataset (typically AgMERRA); [*variable*] is the output variable (of those in Table 2.2); [*crop*] is the crop abbreviation ("mai" for maize, "ric" for rice, "soy" for soybean, "swh" for spring wheat, and "wwh" for winter wheat); and [$start - year$] and [$end - year$] specify the first and last years recorded on file. [C], [T], [W], [N] and [A] indicate the CTWN-A settings, each represented with the respective uppercase letter and the number indicating the level (e.g. "C360_T0_W0_N200" see Table 2.1). Except for the CTWN-A letters, the entire file name

Table 2.2: Output variables, naming convention, and units in the GGCMI Phase 2 protocol. Items in **bold** are the mandatory minimum requirements (if model capacities allow). Other variables are optionally provided depending on availability. Participating modeling groups provided between 3 (PEPIC) and 18 (APSIM-UGOE) of these optional variables.

Variable description	**Variable name**	**Units**
Yield	$yield_ < crop >$	t ha^{-1} yr^{-1} (dry)
Total above ground biomass yield	$biom_ < crop >$	t ha^{-1} yr^{-1} (dry)
Actual planting date	$plant - day_ < crop >$	day of year
Anthesis date †	$anth - day_ < crop >$	days from plant
Maturity date	$maty - day_ < crop >$	days from plant
Applied irrigation water	$pirrww_ < crop >$	mm yr^{-1}
Evapotranspiration (grow. season Σ)	$etransp_ < crop >$	mm yr^{-1} (W_∞)
Transpiration (growing season Σ)	$transp_ < crop >$	mm yr^{-1}
Evaporation (growing season Σ)	$evap_ < crop >$	mm yr^{-1}
Runoff (growing season Σ, sub+surface)	$runoff_ < crop >$	mm yr^{-1}
Total available soil moisture in root zone *	$trzpah2o_ < crop >$	mm yr^{-1}
Total root biomass	$rootm_ < crop >$	t ha^{-1} yr^{-1} (dry)
Total Reactive Nitrogen (Nr) uptake	$tnrup_ < crop >$	kg ha^{-1} yr^{-1}
Total Nr inputs (growing season Σ)	$tnrin_ < crop >$	kg ha^{-1} yr^{-1}
Total Nr losses (growing season Σ)	$tnrloss_ < crop >$	kg ha^{-1} yr^{-1}
Gross primary production (GPP)	$gpp_ < crop >$	gC m^{-2} yr^{-1}
Net primary production (NPP)	$npp_ < crop >$	gC m^{-2} yr^{-1}
[CO_2] response scaler on NPP	$CO_2npp_ < crop >$	- {0..inf}
Water response scaler on NPP	$h2onpp_ < crop >$	- {0..1}
Temperature response scaler on NPP	$tnpp_ < crop >$	- {0..1}
Nr response scaler on NPP	$nrnpp_ < crop >$	- {0..1}
Other nutrient response scaler on NPP	$ornpp_ < crop >$	- {0..1}
[CO_2] response scaler on transpiration	$CO_2trans_ < crop >$	- {0..1}
Maximum stress response scaler	$maxstress_ < crop >$	- {0..1}
Maximum Leaf Area Index (LAI)	$laimax_ < crop >$	m^2 m^{-2}

* growing season Σ, basis for computing average soil moisture

† provided where possible

needs to be in small caps. All filenames include the identifiers *global* and *annual* to distinguish them as global, annual model output, following the updated ISIMIP file naming convention [50].

Output data is provided on a regular geographic grid, identical for all models. Grid cell centers span latitudes -89.75 to 89.75° and longitudes from -179.75 to 179.75°. Missing values where no crop growth has been simulated are distinguished from crop failures: a crop failure is reported as zero yield but non-simulated areas (including ocean grid cells) have yields reported as "missing values" (defined as 1.e+20 in the NetCDF files). Following NetCDF standards, latitude, longitude and time are included as separate variables in ascending order, with units "degrees north", "degrees east", and "growing seasons since 1980-01-01 00:00:00".

Following GGCMI Phase 1 standards, the first entry in each file describes the first complete cropping cycle simulated from the given climate input timeseries. In the AgMERRA timeseries used for GGCMI Phase 2, the first year provided is 1980 but the date of the first entry can vary by crop and location. In the northern hemisphere, for summer crops like maize (sown in spring 1980 and harvested in fall 1980), the first harvest record would be of 1980, but for winter wheat (sown in fall 1980 and harvested in spring 1981) the first harvest record would be of 1981. Output files report the sequence of growing periods rather than calendar years. While there is generally one sowing event per calendar year (since simulations with harmonized growing seasons do not permit double-cropping), in some cases harvest events may skip or repeat within a calendar year. For example, because soybeans in North Carolina are typically harvested well into December, some calendar years may include no harvest (if it is not completed until after Dec. 31) or two harvests (one in January and one 11 months later in the following December).

2.3 Models Contributing

The 12 models participating in GGCMI Phase 2 are listed in Table 2.3. Models differ substantially in structure and parameterization and can be separated into two broad categories: site-based (field-scale) models, and global ecosystem models. The 6 site-based models are APSIM, pDSSAT, and the EPIC family of models; the 6 ecosystem models are LPJmL, LPJ-GUESS, PROMET, CARAIB, ORCHIDEE, and JULES. Models employ a variety of approaches for the core modules such as primary production or evapotranspiration. For primary production, site-based models employ light use efficiency approaches and ecosystem models use photosynthesis approaches. For evapotranspiration, most models use Priestley-Taylor, Penman-Monteith or Hargreaves schemes, but JULES and PROMET utilize a land surface model approach instead. Note that models that share a common genealogy may still use different schemes for evapotranspiration: for example, EPIC-TAMU uses Penman-Monteith and EPIC-IIASA uses Hargreaves. To describe soils, most models use either the Harmonized World Soil Database (HWSD) from the FAO [38] or the ISRIC-WISE database [12] or a derivation thereof. Supplemental Table S1 provides details on these model characteristics as well as on implementation, including spin-up, calibration other than growing season, residue management, and irrigation rules,

Table 2.3 also describes the simulation output contribution of each model to the GGCMI Phase 2 archive. Not all modeling groups provided simulations for the full protocol described above. Given the substantial computational requirements, different participation tiers were specified to allow submission of smaller sub-sets of the full protocol. These subsets were designed as alternate samples across the 4 dimensions of the CTWN space,with *full* (12) and *low* (4) options for the C $\cdot$ N variables, and *full* (63), *reduced* (31), and *minimum* (9) options for T $\cdot$ W variables (described below). All participating modeling groups provided identical coverage of the CTWN parameter space for different crops, but most differed in CTWN coverage of A0 and A1 scenarios. Since the adaptation dimension was defined as

Table 2.3: Models included in GGCMI Phase 2 and the number of CTWN-A simulations performed. The maximum number is 756 for A0 (no adaptation) experiments, and 648 for A1 (maintaining growing length) experiments, since T0 is not simulated under A1. "N" indicates whether the models are able to represent varying nitrogen levels. Each model provides the same set of CTWN simulations across all its modeled crops, but some models omit individual crops. (For example, APSIM-UGOE does not simulate winter wheat.)

Model (Key Citations)	**Maize**	**Soyb.**	**Rice**	**W. wheat**	**S. wheat**	**N**	**Sims per crop (A0 / A1)**
APSIM-UGOE, [96, 75]	X	X	X	–	X	X	44 / 36
CARAIB, [33, 143]	X	X	X	X	X	–	**252 / 216**
EPIC-IIASA, [10]	X	X	X	X	X	X	39 / 0
EPIC-TAMU, [86]	X	X	X	X	X	X	**756 / 648**
JULES, [139, 211, 212]	X	X	X	–	X	–	**252 / 0**
GEPIC, [109, 42]	X	X	X	X	X	X	430 / 181
LPJ-GUESS, [107, 135]	X	–	–	X	X	X	**756 / 648**
LPJmL, [203]	X	X	X	X	X	X	**756 / 648**
ORCHIDEE-crop, [214]	X	–	X	X	–	X	33 / 0
pDSSAT, [36, 94]	X	X	X	X	X	X	**756 / 648**
PEPIC, [110, 111]	X	X	X	X	X	X	149 / 121
PROMET, [59, 117]	X	X	X	X	X	X	261 / 232
Totals	12	10	11	10	11	10	5240 / 3378

a secondary priority for GGCMI Phase 2, some models provided a more limited set of A1 scenarios. Of these, EPIC-IIASA, JULES, and ORCHIDEE-crop provided no A1 scenarios.

The different participation levels are defined by combining the CxN sets with the TxW sets:

- **full**: all 756 A0 simulations (all 12 CxN * all 63 TxW)
- **high**: 362 simulations (all 12 CxN combinations · *reduced* TxW set of 31 combinations)
- **mid**: 124 simulations (*low* 4 CxN combinations · *reduced* TxW set of 31 combinations)

- **low**: 36 simulations (*low* 4 CxN combinations · *minimum* TxW set of 9 combinations)

Of the 12 models submitting data, 6 followed the *full* protocol; these are marked with bold text in the last column of Table 2.3. However, note that two of these models (CARAIB and JULES) cannot represent nitrogen effects explicitly and so do not sample over the nitrogen dimension. Two models followed *high* with minor modifications (GEPIC adding an additional T level and PROMET omitting the intermediate N level). One model (PEPIC) followed *mid* but included an additional C level. Three models approximately followed *low* with APSIM-UGOE and EPIC-IIASA providing some additional TxW levels and ORCHIDEE-crop omitting some TxW combinations.

The combinations of perturbation values in the CxN and TxW parameter spaces used in the various participation levels are chosen to provide maximum coverage over plausible future values. For the CxN space, we specify:

- *full* as 12 pairs, with 4 C values (360, 510, 660, 810 ppm) and 3 N (10, 60, 200 kg ha^{-1} yr^{-1})

- *low* as only 4 pairs: C360_N10, C360_N200, C660_N60, C810_N200

For the TxW space we specify:

- *full* as all 7 T levels and 9W levels.

- *reduced* as 31 alternating combinations, with different Ws for even Ts than for odd Ts. For even Ts (i.e. T0,T2,T4,T6), we use W = -50,-20,0,+30 = 4·4 = 16 pairs. For odd Ts (i.e. T-1,T1,T3) , we use W = -30, -10, +10, +30, inf = 3·5 = 15 pairs.

- *minimum* as 9 combinations: T-1W-10, T0W10, T1W-30, T2W-50, T2W20, T3W30, T4W0, T4Winf, T6W-20

2.3.1 Results

To illustrate the properties of the GGCMI Phase 2 model simulations, we provide an evaluation of model performance by comparing model and historical yields, and show example results that demonstrate the spread of model responses to climate and management inputs.

2.3.2 Evaluation of Model Performance

All models participating in GGCMI Phase 2 have be evaluated against historical yields and site specific experimental data. Most models (9 of 12, all but CARAIB, JULES, and PROMET) have been evaluated in their global setup in the GGCMI Phase 1 evaluation exercise [126], and many have used the GGCMI Phase 1 online tool to similarly evaluate subsequent model versions [e.g. 203]. Evaluating the performance of crop models in the GGCMI Phase 2 archive is complicated by the artificial nature of the protocol: the settings in the CTWN-A experiment design do not reflect actual conditions in the real world. The protocol includes one scenario of near-historical climate inputs (T_0, W_0, C_{360}), but the prescribed uniform nitrogen application levels do not reflect real-world fertilizer practices. Models also omit detailed calibrations to reflect the performance of historical cultivars.

We provide a partial evaluation of the models' skill in reproducing crop yield characteristics using the methodology of [126], developed for GGCMI Phase 1. [126] evaluate how well model crop yield responses in a historical run capture real-world yield variations driven by year-to-year temperature and precipitation variations. Following this approach, we compare yields in the GGCMI Phase 2 baseline simulations with detrended historical yields from the Food and Agriculture Organization of the United Nations [46] by calculating the Pearson product moment correlation coefficient over 26 years of yield. The procedure is sensitive to the detrending method and the area mask used to aggregate yields; we use a 5-year running mean removal and the MIRCA2000 cultivation area mask for aggregation. In some cases the model timeseries are shifted by one year to account for discrepancies in FAO or model year

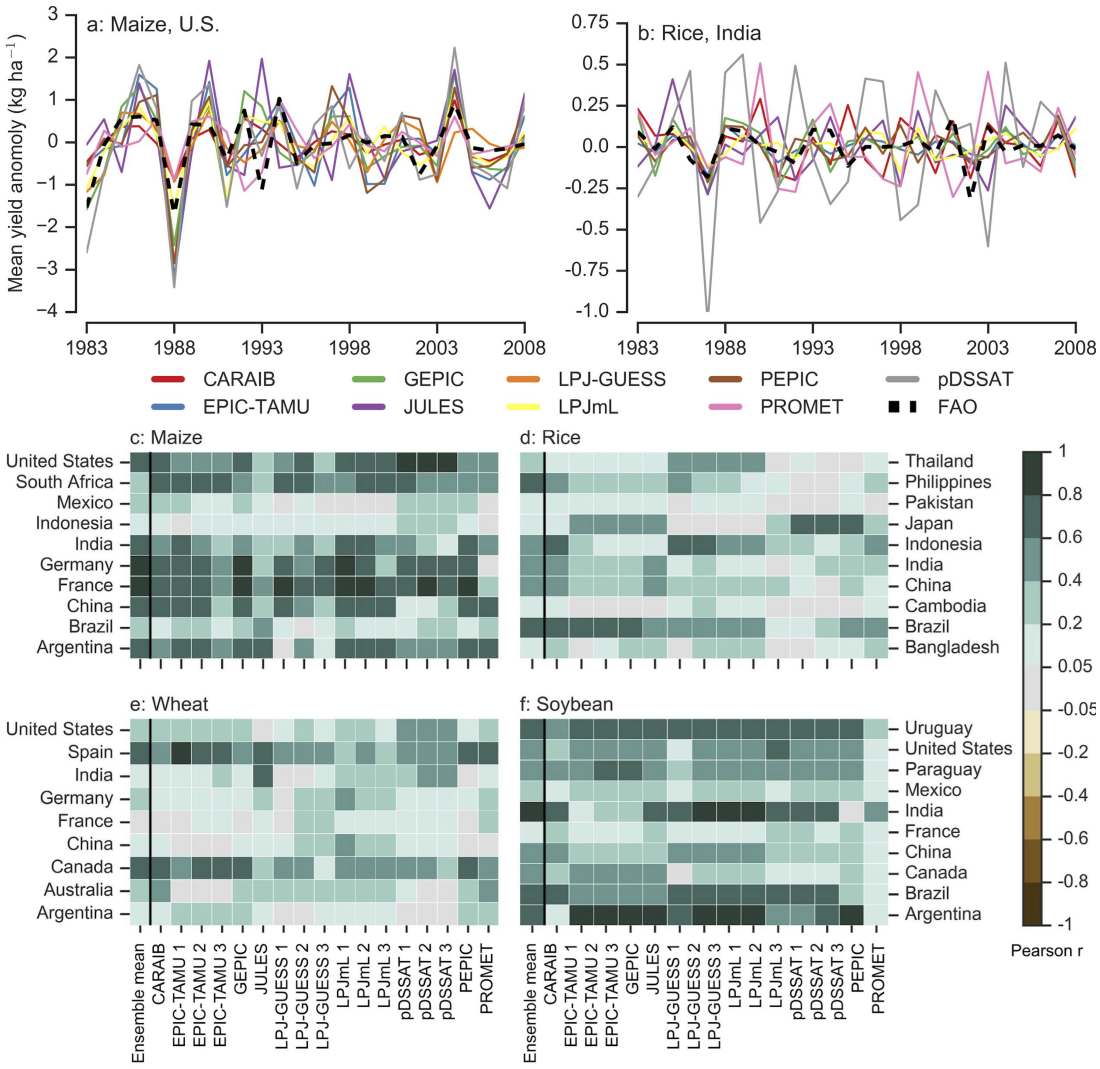


Figure 2.2: Assessment of crop model performance in GGCMI Phase 2, following the protocol of GGCMI Phase 1 [126]. **Top:** example timeseries comparison between simulated crop yield and FAO country statistics [46] at the country level for two example high production countries: US maize, and rice in India, both for the 200 kg ha^{-1} nitrogen application level. **Bottom:** heatmaps illustrating the Pearson r correlation coefficient between the detrended simulated and observed country-level mean yields for the top 10 countries by production for each crop, of those countries with continuous FAO data over 1981-2010. We show separate comparisons for simulations with the three different nitrogen application levels, denoted 1, 2, 3 for 10, 60, and 200 kg N ha^{-1}, respectively. Left column shows correlation of ensemble mean yields with FAO data. Because FAO does not distinguish between wheat types, we sum simulated spring and winter wheat for models that provide both (See Table 2.3.). Note that differences by region and crop are stronger than difference between models, e.g. horizontal bars are more similar in color than vertical bars. Source: [47].

reporting. Because the GGCMI Phase 2 protocol imposes fixed, uniform nitrogen application levels that are not realistic for individual countries, we evaluate control runs for each model at multiple N levels whenever possible. Nine of the GGCMI Phase 2 models (Table 3) provide historical runs for all three nitrogen levels (10, 60, and. 200 kg $ha^{-}1$ $yr^{-}1$).

As expected due to the unrealistic features described above, correlation coefficients for the GGCMI Phase 2 simulations are slightly lower than those found in the Phase 1 evaluation, but models show reasonable fidelity at capturing year-over-year variation (Figure 2.2). For example, global correlation coefficients in Phase 1 and Phase 2, respectively, are 0.89 and 0.74 for maize, 0.67 and 0.64 for wheat, and 0.64 and 0.59 for soybeans. (Phase 1 values are from Figures 1–4 and 6 in [126].) Differences in fidelity between regions and crops exceed differences between models: that is, Figure 2.2(c)–2.2(f) shows more color similarity in horizontal than vertical bars. For example, maize in the United States is consistently well-simulated while maize in Indonesia is problematic (mean Pearson correlation coefficients of 0.68 and 0.18, respectively). Note that in this methodology, simulations of crops with low year-to-year variability such as irrigated rice and wheat will tend to score more poorly than those with higher variability. In some cases, especially in the developing world, low correlation coefficients may point to reporting problems in the FAO statistics and to real-world variability caused by variations in management rather than weather [157, 126]. No single model consistently exhibits greater fidelity than others. Instead, each model shows near best-in-class performance for at least one location-crop combination. For example, pDSSAT is the best model for maize in the US, LPJmL and GEPIC are best in Germany, PROMET is best in Argentina, and PEPIC and LPJ-GUESS are best in France.

2.3.3 Model Crop Yield Responses Under CTWN Forcing

Crop models in the GGCMI Phase 2 ensemble show broadly consistent responses to climate and management perturbations in most regions, with a strong negative impact of increased

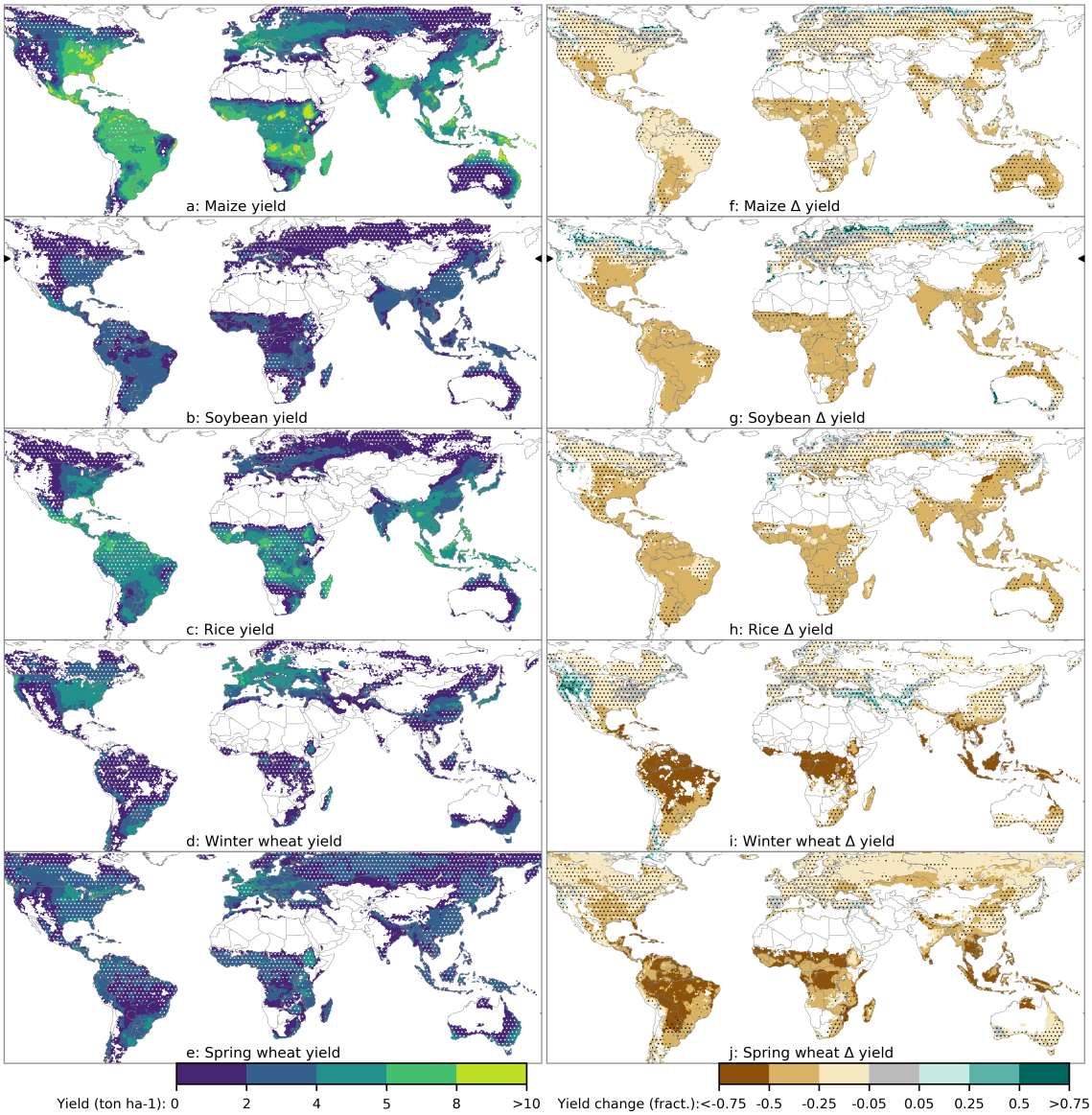


Figure 2.3: Illustration of the spatial patterns of baseline yields (left) and yield changes (right) in the GGCMI Phase 2 ensemble. Left column shows multi-model climatological(30 year) median yields for the baseline scenario, with white stippling indicating areas where these crops are not currently cultivated. Areas with less that 0.5 ton ha^{-1} in the baseline are masked. Absence of cultivation aligns well with the lowest yield contour (0-2 ton ha^{-1}). Right column shows multi-model mean fractional yield changes in the T+4 °C scenario relative to the baseline scenario. Areas without stippling are those where models agree on changes: the multi-model mean fractional change exceeds the standard deviation of changes in individual models. Stippling indicates areas of low confidence ($\Delta < 1\sigma$). Some spatial structure in projected changes at high latitudes may be due to differences in model coverage; see Figure 2.1. Source: [47].

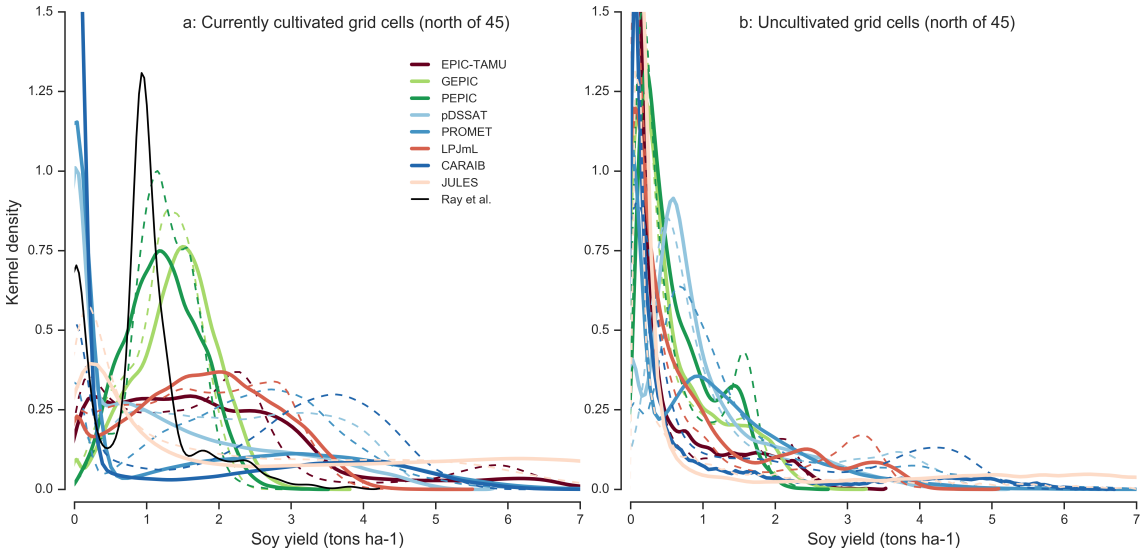


Figure 2.4: Model probability densities for soybean yields at latitudes north of 45° in historical and warming simulations in the A0 case. While 10 GGCMI Phase 2 models provide simulations (Table 2.3); we show 8 representative models for clarity. Probability density functions are estimated separately for locations with some current cultivation (left, approximately 2500 grid cells, unweighted by cultivated area) and for uncultivated locations (right, approximately 1500 grid cells), for baseline historical (solid) and T+4 (°C) (dashed) simulations. Black line in left panel shows actual yields from 1997-2003 derived from [157]. For historical simulations, models agree on low potential yields in currently uncultivated areas (right) but disagree widely on yields in currently cultivated areas (left). Color code groups models into those with realistic yield distributions peaking at 1-2 ton ha^{-1} (green), those with flatter distributions extending to unrealistically high values (red), and those with predominantly zero yields (blue). "Green" models show slight decreases under T+4 warming, "red" models moderate increases, and "blue" models large increases. Source: [47].

temperature in all but the coldest regions. Mapping the distribution of baseline yields and yield changes shows the geographic dependencies that underlie these results. Absolute yield potentials show strong spatial variation, with much of the Earth's surface area unsuitable for any of these crops (Figure 2.3, left). Crop yield changes under climate perturbations also show distinct geographic patterns (Figure 2.3, right, which shows fractional yield differences between the baseline and T+4 A0 scenarios). In general, models agree most on yield response in regions where yield potentials are currently high and therefore where crops are currently grown. In A0 simulations, models show robust decreases in yields at low latitudes, and highly uncertain ensemble mean increases at most high latitudes. Low latitude yield reductions are due in part to shortening of the growing season under warming and in part to the direct

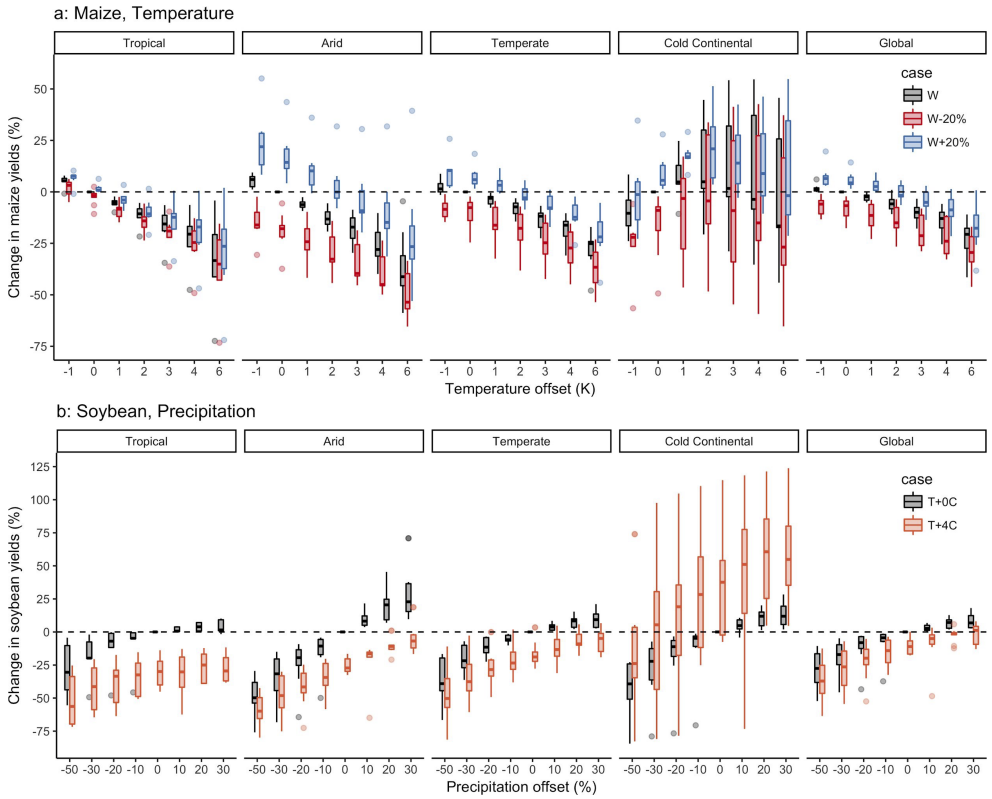


Figure 2.5: Illustration of the distribution of regional yield changes across the multi-model ensemble, split by Köppen-Geiger climate regions, and with global response in rightmost panel. Y-axis is the fractional change in the regional average climatological (30-year mean) potential yield relative to the baseline. Box-and-whiskers plots show distribution across models, with median marked; edges are first and third quartiles and whiskers extend to 1.5·IQR. Figure shows all simulated grid cells for each model; see Supplemental Figure S10-S13 for only currently-cultivated land. We highlight responses to individual factors; note that results are not directly comparable to simulations of realistic projected climate scenarios with identical global mean changes. Models generally agree outside high-latitude regions, with projected changes exceeding inter-model variance. **Top:** Response of rainfed maize to applied uniform temperature perturbations, for three discrete precipitation perturbation levels (-20%, 0%, and +20%), with [CO_2] and nitrogen held constant at baseline values (360 pmm and 200 kg ha^{-1} yr^{-1}). Outliers in the tropics (strong negative impact of higher T) are the pDSSAT model; outliers in the arid region (strong positive impact of higher P) are JULES. **Bottom:** Response of rainfed soybeans to applied uniform precipitation perturbations, for two discrete temperature levels. Cases with reduced precipitation show greater inter-model spread than those with increased precipitation. At very large precipitation increases, yield changes level out: benefits saturate once water availability is no longer limiting. Precipitation changes are more important in the arid region, as expected. Note the large uncertainty in the cold continental region, also illustrated in Figures 2.3 and 2.4. Source: [47].

effects of higher temperature. In A1 simulations, where growing seasons length does not change, temperature-related reductions in yield are more muted (see Supplemental Figure

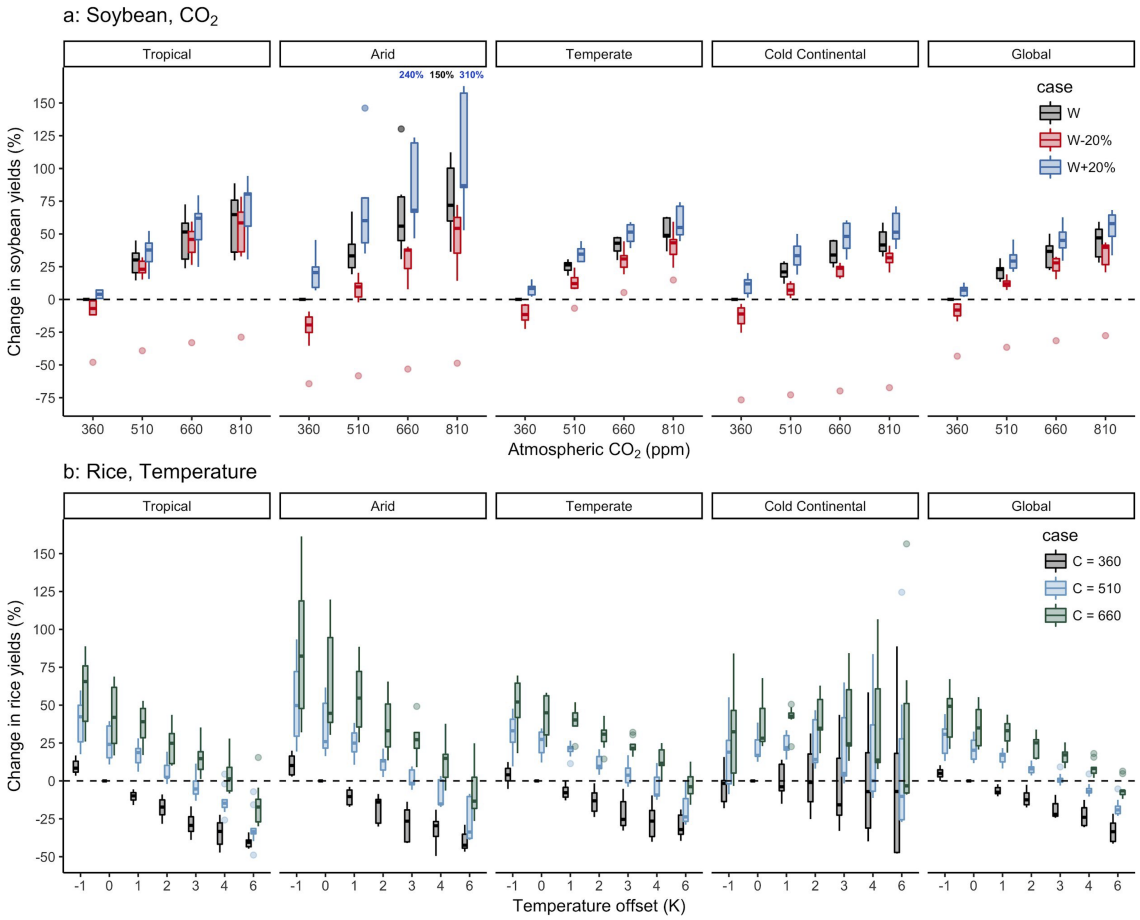


Figure 2.6: Illustration of the distribution of regional yield changes across the multi-model ensemble, here for soybeans and rice for the A0 case. Conventions as in Figure 2.5. **Top:** Response of rainfed soybeans to atmospheric CO_2, for three discrete precipitation perturbation levels with temperature and nitrogen held constant at baseline values. Low outliers are the EPIC-TAMU model and the high outliers in the Arid region are the JULES model. Reduced precipitation tends to steepen the [CO_2] response and increased precipitation tends to flatten it, as expected. Reduced precipitation tends to increase the inter model spread, especially at the highest [CO_2] levels. **Bottom:** Response of irrigated rice for three discrete [CO_2] levels, with nitrogen and precipitation held constant. [CO_2] does not change the nature of temperature response respective to baseline as the slopes at each [CO_2] level are relatively constant. Source: [47].

S14). In both A0 and A1 simulations, models show some increases in high mountain regions that are currently cold-limited.

Projections of strong yield growth at higher latitudes should be treated with caution, since the effects evident in Figure 2.3 are due in part to inaccuracies in model representations of present-day crop yields. For example, at latitudes north of 45°, the GGCMI Phase 2

models collectively suggest strong (but uncertain) growth in soybean yields under warmer conditions (Figure 2.3, g). However, model differences are greater in the baseline than future simulations, and greatest in currently-cultivated areas (Figure 2.4). Both the mean projected growth and the inter-model spread are driven by three models that show almost zero present-day potential soybean yields across the entire high-latitude region, even in locations where soybeans are currently grown (Figure 2.4, left). PROMET, for example, involves a stronger response to cold than other models (e.g. LPJmL) with frost below -8 °C irreversibly killing non-winter crops and prolonged periods of below-optimum temperatures also leading to complete crop failure. Over the high-latitude regions simulated by both models, 52% of grid cells in PROMET report 0 yield in the present climate vs. 11% of cells in the T+4 scenario, leading to a strong yield gain in warmer future climates. In LPJmL outputs, the same high-latitude area is deemed suitable for cultivation even in baseline climate, with crop failure rates of 4% and 5% in present and T+4 cases, so that projected yield changes are modest (Figure 2.4). These spurious low baseline yields result in very large fractional changes in the T+4 warming scenario, when all models agree that conditions become favorable for soybeans. Those models that most accurately reproduce present-day high-latitude soybean yields of 1-2 ton ha^{-1} [157] in fact show a slight decrease in yield under a warming scenario (Figure 2.4, left). Apparent future yield increases in the multi-model mean are driven by the least realistic simulations.

The GGCMI Phase 2 exercise offers the opportunity to examine and characterize not just crop response to a single temperature change but nonlinearities in responses and interactions between factors. We illustrate a few of these relationships in Figures 2.5-2.6 using A0 simulations to capture maximum climate effects. We choose crops and factors whose effects are reasonably well understood, and show that these are reproduced in models. It is expected, for instance, that increases in precipitation should buffer the effects of warmer temperatures and that [CO_2] increases should reduce damage to crops in scenarios where water is lim-

ited. Models generally confirm expected behavior but also provide insight into unforeseen interactions. To show geographic effects, we divide model responses in Figures 2.5-2.6 by the primary Köppen-Geiger climate regions [175], showing the yield changes across all simulated grid cells in each region. In each panel we examine relationships between two factors, showing yield response against one for several scenarios of the other, in box plots that show the inter-model spread. The responses highlighted here are qualitatively similar across all crops included in this study (Supplemental Figures S5–S9 for all simulated area and S10–S13 for cultivated area only). For all crops, warming scenarios with precipitation held constant produce yield decreases in most regions. These impacts are robust for even moderate climate perturbations. For rainfed maize, even a 1°C temperature increase with other factors held constant produces a median regional decline in potential yield that exceeds the variance across models, in all but the "cold-continental" regions (Figure 2.5a). The remaining areas ("warm temperate", "equatorial", and "arid") account for nearly three-quarters of global maize production. In the high-latitude "cold-continental" region, potential yield changes are positive but highly uncertain, for the reasons discussed previously; uncertainties are larger even for maize than for soybeans. (Compare Figures 2.5a and 2.5b.) Temperature effects are somewhat nonlinear, with the largest impacts for maize in the warm "tropical" region (for soybeans, temperature effects are more complex; see Supplemental Figure S5). Precipitation effects on rainfed crops are more strongly nonlinear. The curvature of the precipitation response can be seen by eye in Figure 2.5b: soybean yields are strongly negatively impacted by reduced rainfall, peak under increased precipitation of 20%, and actually decline at higher precipitation levels.

As expected, precipitation and temperature effects interact, with increases in precipitation buffering yield responses to temperature. Increased rainfall mitigates the negative impacts of warmer temperatures caused by increased evapotranspiration [e.g. 219]. For maize, the effect is relatively modest outside the "arid" regions (Figure 2.5a). Globally,

a 4°C temperature rise with no change in precipitation results in median loss of ∼13% of rainfed maize, with all models showing a negative response. With a 20% increase in precipitation, the median yield loss is ∼8%. For soybeans, the equivalent values are ∼11% and 1%, respectively. Decreased rainfall, on the other hand, amplifies yield losses and also increases inter-model variance. That is, models agree that the response to decreased water availability is negative in sign but disagree on its magnitude. Outside of arid regions, the interaction effect itself shows little nonlinearity (i.e. response slopes in Figures 2.5a and 2.5b are roughly parallel). As expected, irrigated crops are more resilient to temperature increases in all regions, especially so where water is already limiting (other than winter wheat, Supplemental Figure S9).

Increased [CO_2] boosts yields overall through the well-known [CO_2] fertilization effect (Figure 2.6). The effect is strongest for the C3 crops (wheat, soybeans, and rice), while maize, a C4 grass, has a comparatively muted response. We show irrigated rice and rainfed soy in Figure 2.6 as representative C3 crops. The effect of [CO_2] on yields is nonlinear, as expected, with significant benefit from small increases but with effects plateauing at higher concentrations (Figure 2.6). [CO_2]and temperature effects show minimal interaction. This effect is seen in Figure 2.6a, which sho ws nearly parallel response slopes at differen t[CO_2] levels. That is, [CO_2] fertilization does little to change the nature of the temperature response. On the other hand, [CO_2] and precipitation effects interact strongly, as expected since higher[CO_2]levels allow reduced stomatal conductance and evapotranspiration losses, mitigating the effect of reduced rainfall [e.g. 119]. This interaction is seen in Figure 2.6b as smaller yield losses from reduced rainfall when [CO_2] levels are higher. For example, for soy, raising [CO_2] to 510 ppm actually outweighs the multi-model median damages caused by a 20% precipitation reduction in all climate regions. All crops show similar behavior, but note that model uncertainties for wheat are substantially higher than those for other crops. (Compare Figure 2.6a for soy and Supplemental Figure S7 for wheat). We show some addi-

tional cases in Supplemental Material. As noted previously, the A1 adaptation simulations involve significantly moderated temperature impacts relative to the A0 simulations shown here (Supplemental Figure S14). Supplemental Figures S15 and S16 show the response in the nitrogen dimension and an irrigation water demand response example.

2.4 Discussion and Conclusions

The GGCMI Phase 2 experiment provides a database designed to allow detailed study of crop yields in process-based models under climate change. While previous crop model inter-comparison projects in the climate change context have focused on simulations along realistic projected climate scenarios [e.g. 165], the use of systematic input parameter variations in GGCMI Phase 2, with up to 756 scenarios, allows not only comparing yield sensitivities to changing climate and management inputs but also evaluating the complex interactions between important driving factors: [CO_2], temperature, water supply, and applied nitrogen. The global extent of the experiment also allows identifying geographic shifts in high potential yield locations. With 12 participating models and 31 simulation years per scenario, the complete database constitutes over 150,000 years of gridded global yield simulation output for each crop.

Preliminary results shown here highlight some of the insights facilitated by the simulation exercise and lend confidence in the models. In validation tests of historical simulations, year-over-year correlations in modeled and actual country-level yields are similar to those of GGCMI Phase 1. In simulations of scenarios with perturbed climate and management factors, models broadly agree on changes outside the high latitudes, with the magnitude of changes at nearly all perturbation levels exceeding the inter-model spread. At high latitudes, differences between models may result from differences in their assumed yields in current cold conditions. In simulations with multiple perturbations, interactions between major yield drivers (e.g. temperature and precipitation in Figure 2.5, or precipitation and [CO_2]

in Figure 2.6) generally follow expectations and produce physically reasonable responses in crop yields.

Users should however be aware of some limitations of the GGCMI Phase 2 experiment that affect its potential applications. First, absolute yield values in the baseline scenario, driven by 1981–2010 historical climate, will generally not match observed yields over this time period. In order to match current yields, process-based models must be re-tuned to account for the constant evolution of crop cultivar genetics and management practice [e.g. 93]. GGCMI Phase 2 is intended as a study of model-projected changes under future climate change, which are assumed insensitivee to the adjustments needed to reproduce present-day yields. The baseline scenario also includes no trend in $[CO_2]$, and no individual case involves realistic country-specific nitrogen application levels [34].

The second major caveat is that no individual GGCMI Phase 2 simulation is itself a realistic future yield projection. The uniform applied offsets in temperature and precipitation sample over potential changes, and do not individually capture the spatially heterogeneous warming and precipitation changes expected in realistic climate projections. GGCMI Phase 2 simulation results can be used for impacts projection, but only with the construction of an emulator of crop yield response to climatological changes, which can then by driven by arbitrary climate scenarios. Such emulators are shown to accurately reproduce crop model output under realistic climate projections, even though the GGCMI Phase 2 experiment does not sample over potential changes in the higher-order moments in temperature and precipitation distributions. Note that some factors that may affect future climate-driven yield impacts cannot be captured by the GGCMI Phase 2 models in any usage, since models do not include representations of pests, diseases, and weeds. Off-line crop model simulations can also not capture any feedbacks on the climate from land use, such as irrigation impacts on humidity [e.g. 29].

We expect that the GGCMI Phase 2 simulations will yield multiple insights in future

studies. Potential applications include, as mentioned, the construction of emulators and yield response surfaces that can be used for both model diagnosis and impacts assessment. Specific studies could include analyzing the drivers of temperature-related yield losses (which may be due to both direct thermal effects or to shortening growing seasons); the benefits of adaption; interactions between [CO_2] and water or other CTWN factors affecting yield; changes in nitrogen use efficiency; geographic shifts in regional production; regional differences in yield sensitivities to CTWN-A factors. Emulators based on the dataset can be used to identify hotspots of crop system vulnerability, and to conduct rapid assessment of new climate projections. In general, the development of multi-model ensembles involving systematic parameter sweeps has large promise both for increasing understanding of potential future crop responses and for improving process-based crop models.

2.5 Data Availability

The simulation outputs of the mandatory 7 output variables (Table 2.2) are available on zenodo.org. See Appendix 2.5 for data DOIs. All other simulation output variables are available upon request to the corresponding author. The scripts for generating the spring wheat and winter wheat growing seasons and second fertilizer dates and the quality screening script is available at *https://github.com/RDCEP/ggcmi/blob/phase2/*. All input data are available via globus.org (registration required, free of charge): Minimum cropland mask is available at *tinytextit.com/5hd72bk7* choose the file boolean_cropmask_ggcmi_phase2.nc4 Growing period data for wheat is now divided up into winter and spring wheat, available at *tinytextit.com/5hd72bk7* whereas all other growing season data (maize, rice, soybean) are the same as in Phase 1 (version 1.25), available at *https://tinytextit.com/urzjjrnc* Simulation data are published in crop- and GGCM-specific packages on *https://zenodo.org/*, in order to break down the overall data amount into manageable packages (50GB per archive).

CHAPTER 3

THE GGCMI PHASE 2 EMULATORS: GLOBAL GRIDDED CROP MODEL RESPONSES TO CHANGES IN CO_2, TEMPERATURE, WATER, AND NITROGEN

Published in *Geoscientific Model Development* in collaboration with:

Christoph Müller[3], Joshua Elliott[2,4], Alex C. Ruane[5], Jonas Jägermeyr[4,2,3,5], Abigail Snyder[6], Marie Dury[7], Pete D. Falloon[8], Christian Folberth[9], Louis François[7], Tobias Hank[10], R. Cesar Izaurralde[11,12], Ingrid Jacquemin[7], Curtis Jones[11], Michelle Li[2,13], Wenfeng Liu[14,15], Stefan Olin[16], Meridel Phillps[5,17], Thomas A. M. Pugh[18,19], Ashwan Reddy[11], Karina Williams[8,20], Ziwei Wang[1,2], Florian Zabel[10], and Elisabeth J. Moyer[1,2]

[1]Department of the Geophysical Sciences, University of Chicago, Chicago, IL, USA [2]Center for Robust Decision-making on Climate and Energy Policy (RDCEP), University of Chicago, Chicago, IL, USA [3]Potsdam Institute for Climate Impact Research, Member of the Leibniz Association, Potsdam, Germany [4]Department of Computer Science, University of Chicago, Chicago, IL, USA [5]NASA Goddard Institute for Space Studies, New York, NY, United States [6]Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD, USA [7]Unité de Modélisation du Climat et des Cycles Biogéochimiques, UR SPHERES, University of Liège, Belgium [8]Met Office Hadley Centre, Exeter, United Kingdom [9]Ecosystem Services and Management Program, International Institute for Applied Systems Analysis, Laxenburg, Austria [10]Department of Geography, Ludwig-Maximilians-Universität, Munich, Germany [11]Department of Geographical Sciences, University of Maryland, College Park, MD, USA [12]Texas Agrilife Research and Extension, Texas A&M University, Temple, TX, USA [13]Department of Statistics, University of Chicago, Chicago, IL, USA [14]EAWAG, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland [15]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, France. [16]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden [17]Earth Institute Center for Climate Systems Research, Columbia University, New York, NY, USA [18]School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK. [19]Birmingham Institute of Forest Research, University of Birmingham, Birmingham, UK. [20]Global Systems Institute, University of Exeter, Laver Building, North Park Road, Exeter, EX4 4QE, UK

Statistical emulation allows combining advantageous features of statistical and process-based crop models for understanding the effects of future climate changes on crop yields. We describe here the development of emulators for nine process-based crop models and five crops using output from the Global Gridded Model Intercomparison Project (GGCMI) Phase 2. The GGCMI Phase 2 experiment is designed with the explicit goal of producing a structured

training dataset for emulator development that samples across four dimensions relevant to crop yields: atmospheric carbon dioxide (CO_2) concentrations, temperature, water supply, and nitrogen inputs (CTWN). Simulations are run under two different adaptation assumptions: that growing seasons shorten in warmer climates, and that cultivar choice allows growing seasons to remain fixed. The dataset allows emulating the climatological mean yield response of all models with a simple polynomial in mean growing-season values. Climatological mean yields are a central metric in climate change impact analysis; we show here that they can be captured without relying on interannual variations. In general, emulation errors are negligible relative to differences across crop models or even across climate model scenarios; errors become significant only in some marginal lands where crops are not currently grown. We demonstrate that the resulting GGCMI emulators can reproduce yields under realistic future climate simulations, even though the GGCMI Phase 2 dataset is constructed with uniform CTWN offsets, suggesting that the effects of changes in temperature and precipitation distributions are small relative to those of changing means. The resulting emulators therefore capture relevant crop model responses in a lightweight, computationally tractable form, providing a tool that can facilitate model comparison, diagnosis of interacting factors affecting yields, and integrated assessment of climate impacts.

3.1 Introduction

Improving our understanding of the impacts of future climate change on crop yields is critical for global food security in the twenty-first century. Projections of future yields under climate change are generally made with one of two approaches: either process-based models, which simulate the process of photosynthesis and the biology and phenology of individual crops, or statistical models, which use historical weather and yield data to capture relationships between observed crop yields and major drivers. Process-based crop models provide some advantages, including capturing the direct effects of [CO_2] fertilization and allowing projections

in areas where crops are not currently grown. However, they are computationally expensive, and can be difficult or impossible to directly integrate into integrated climate change impacts assessments. Statistical crop models can only capture crop responses under the range of current conditions but have several advantages: they implicitly include management and behavioral practices that are difficult to model explicitly, and they are typically simple analytical expressions that are easily implemented by downstream impact modelers. Both types of models are routinely used, and comparative studies have concluded that when done carefully, both approaches can provide similar yield estimates [e.g. 114, 124, 161, 220, 108].

Statistical emulation allows combining some of the advantageous features of both statistical and process-based models. The approach involves constructing a "surrogate model" of numerical simulations by using their output as training data for a statistical representation [e.g. 134]. Emulation is particularly useful in cases where simulations are complex and output data volumes are large and has been used in a variety of fields, including hydrology [e.g. 158], engineering [e.g. 191], environmental sciences [e.g. 155], and climate [e.g. 20, 72]. For agricultural impacts studies, emulation of process-based models allows capturing key relationships between input variables in a lightweight, flexible form that is compatible with economic studies. The resultant statistical model can produce yield projections under arbitrary emissions scenarios and is an important diagnostic tool for model comparison and model evaluation.

Interest is rising in applying statistical emulation to crop models, and multiple studies have developed crop model emulators in the past decade. Early studies proposing or describing potential crop yield emulators include [76, 152, 114], and [39]. Studies developing single-model emulators include [74] for the CropSyst model, [170] for the CERES wheat model, and [141] for the LPJmL model. More recently, emulators have begun to be used in the context of multi-model intercomparison, with multiple authors [15, 14, 140, 122] using them to analyze the five crop models of the Inter-Sectoral Impact Model Intercomparison

Project (ISIMIP). ISIMIP offers a relatively large training set – control, historical, and several Representative Concentration Pathway (RCP) scenarios using output from up to five climate models [207, 50] – and choices of emulation strategy differ. [15] and [14] use historical and RPC8.5 scenarios, combine multiple climate model projections for RCP8.5, and regress across soil regions. [140] use global mean temperature change (and $[CO_2]$) as regressors, and then pattern-scales to emulate local yields. [122] compare emulated and observed historical yields, using local weather data and a historical crop simulation. The constraints of the ISIMIP experiment mean that all these efforts do share important common features. All emulate annual crop yields along an entire scenario or scenarios, and all future climate scenarios are non-stationary, with important covariates (temperature and precipitation for example) evolving simultaneously.

An alternative approach to emulation involves construction of a "parameter sweep" training set, a collection of multiple stationary scenarios that systematically cover a range of input parameter values. A parameter sweep offers several important advantages for emulation over an experiment in which climate evolves over time. First, it allows separating the effects of different variables that affect yields but that are highly correlated in realistic future scenarios like those used in ISIMIP (e.g. $[CO_2]$ and temperature). Second, it allows making a distinction between year-to-year yield variations and climatological changes, which may involve different responses to the particular climate regressors used [e.g. 172]. For example, if year-to-year yield variations are driven predominantly by variations in the distribution of temperatures throughout the growing period, and long-term climate changes are driven predominantly by additive mean shifts, then regressing on the mean growing period temperature will produce different yield responses at annual vs. climatological timescales.

Systematic parameter sweeps have begun to be used in crop model evaluation and emulation, with early efforts in 2014 and 2015 [173, 143], and several recent studies in 2018 and 2019 [51, 176, 188]. These three studies sample multiple perturbations to temperature

and precipitation, and two of the three add [CO_2] as well, for a total of 132, 99, and 220 different combinations, respectively. All take advantage of the structured training set to construct emulators ("response surfaces") of climatological mean yields, omitting year-to-year variations. All the 2018–2019 papers have some limitations, however, for assessing global agricultural impacts, including that none evaluate responses in every grid cell globally. Two involve many crop models but only one crop (wheat) [51, 176] and cover only 1-4 individual sites. [188] analyzes five crops over ∼1,000 sites with individual site-specific crop models, and extrapolates in space to estimate mean latitudinal responses.

In this paper we describe a set of globally-gridded crop model emulators developed from the new parameter-sweep dataset of the Global Gridded Crop Model Intercomparison (GGCMI) Phase 2 effort. GGCMI Phase 2, a part of the Agricultural Model Intercomparison and Improvement Project (AgMIP) [163, 165], provides the first near-global-coverage systematic parameter sweep of multi-model crop simulations consisting of up to 756 combinations in [CO_2], temperature, water supply, applied nitrogen, and two different assumptions on growing season adaptation ("A0": none and "A1": retaining growing season length) [CTWN-A, 47, 120]. The experiment is designed to allow diagnosing the impacts on crop yields of both individual factors and their joint effects, and to allow construction of crop model emulators. In Section 3.2 following, we describe the training dataset, including the GGCMI Phase 2 experimental protocol and model participation and the models' differing year-to-year and climatological mean responses. Section 3.3 describes the statistical model used for emulation, Section 3.4 evaluates measures of emulator fidelity, and Section 3.5 shows examples of preliminary results.

Table 3.1: Crop models included in GGCMI Phase 2 emulators and the number of CTWN-A (Carbon dioxide, Temperature, Water, Nitrogen, Adaptation) simulations performed for each model. The maximum number is 756 for A0 (no adaptation) experiments, and 648 for A1 (maintaining growing season length) experiments since T0 is not simulated under A1. "N" indicates whether the models are able to represent varying nitrogen levels. Each model provides the same set of CTWN simulations across all its modeled crops, but some models omit individual crops. Table adapted from [47]. For clarity, three simulation models that submitted data to the GGCMI Phase 2 experiment [47] are not shown here, as they provided a training set too small to be used in emulation.

Model (Key Citations)	**Maize**	**Soyb.**	**Rice**	**W. wheat**	**S. wheat**	**N**	**Sims per crop (A0 / A1)**
CARAIB, [33, 143]	X	X	X	X	X	–	**252 / 216**
EPIC-TAMU, [86]	X	X	X	X	X	X	**756 / 648**
JULES, [139, 211, 212]	X	X	X	–	X	–	**252 / 0**
GEPIC, [109, 42]	X	X	X	X	X	X	430 / 181
LPJ-GUESS, [107, 135]	X	–	–	X	X	X	**756 / 648**
LPJmL, [203]	X	X	X	X	X	X	**756 / 648**
pDSSAT, [36, 94]	X	X	X	X	X	X	**756 / 648**
PEPIC, [110, 111]	X	X	X	X	X	X	149 / 121
PROMET, [59, 117, 217]	X	X	X	X	X	–	261 / 232

3.2 Training Dataset

3.2.1 The GGCMI Phase 2 Dataset

The GGCMI Phase 2 simulations are described in detail in Chapter 2, but we summarize briefly here. The experiment involves 9 different globally gridded crop models, each simulating multiple crops (maize, rice, soybean, and spring and winter wheat) across a systematic parameter sweep of as many as 756 combinations, each driven by a historical climate timeseries with systematic perturbations to [CO_2], temperature, water supply, and nitrogen application (CTWN). The simulation protocol involves 4 levels of atmospheric [CO_2], 7 of temperature, 9 of water supply, and 3 of applied nitrogen, and simulations are repeated for 2 adaptation scenarios: "A0" simulations assume no adaptation in cultivar choice, so that growing seasons shorten in warmer climates, and "A1" simulations assume that adaptation in cultivar choice maintains fixed growing seasons. The complete protocol for each modeling

group involves up to 43,524 years of global simulated output for each crop. Because the computational demand is high, modeling groups were allowed to submit at various specified levels of participation, with the lowest recommended level of participation consisting of 20% of the maximum possible simulations. The mean participation level is 65%, but three models (APSIM-UGOE, EPIC-IIASA, and ORCHIDEE-crop) contributed data below the recommended threshold ($< 5\%$ of the full protocol) and are excluded here since they could not be robustly emulated. Table 3.1 shows the participating models and the number of simulation scenarios that each provides, and Supplemental Figure S17 shows model sampling density. See [47] for the parameter combinations included by each model. Table 2.1 shows the specified input values; we sample across all parameter combinations.

Each individual crop model simulation is run for 31 years over historic weather for the period of 1981-2010, with added uniform perturbations to any of the CTWN variables. Historical weather is taken for most models from the AgMERRA [171] historical daily climate data product, but the PROMET model uses the ERA-Interim reanalysis [30] and the JULES model uses a bias-corrected version of ERA-Interim, WFDEI [WATCH-Forcing-Data-ERA-Interim, 208] as these groups have specific sub-daily input data requirements. Temperature perturbations are applied as additive mean shifts, water supply as fractional multipliers to precipitation (except in the irrigated W_∞ case), and $[CO_2]$ and nitrogen application levels are specified as fixed values. Models provide near-global output at 0.5 degree latitude and longitude resolution for each simulation year, including areas not currently cultivated. Crop models included here are not formally calibrated, given that there is no adequate calibration target for gridded global-scale crop model simulations. This may be a shortcoming if targeting absolute yield levels, but when focusing on relative yield changes, calibration can also have negative effects on model skill [126]. In analyses where we distinguish yields over currently cultivated land, we use the harvested area masks of [145]. (See Supplemental Figure S18 for maps of cultivated area.)

3.2.2 Climatological vs. Year-to-Year Responses

The central metric in assessments of climate change impacts on crop yields is the change in multi-annual means [e.g. 180, 22, 165, 127, 220, 77]. Agricultural impacts assessments work with multi-annual yields, as their analysis frameworks require information on long-term effects [e.g. 132, 189, 210, 64, 188]. Changes in extremes or year-to-year variability are other metrics of potential interest, but are often not explicitly considered in integrated climate change impact assessments or land-use change projections. For this reason we emulate the climatological mean response, i.e. the change in aggregated mean yield in each 30-year simulation. Emulation then becomes relatively straightforward since changes in time-averaged yields are considerably smoother than those in year-to-year yield response. In the GGCMI Phase 2 simulation output dataset, year-to-year responses to weather are also often quantitatively distinct from responses to climatological shifts, with the discrepancy especially strong in wheat and rice. The difference in behavior is illustrated in Figure 3.1, which shows irrigated and rainfed maize and wheat in representative locations. When discrepancies are large, year-to-year responses are generally stronger than climatological ones, but exact responses differ by crop and region and even by model within GGCMI Phase 2.

While differences in responses at different timescales can arise for many reasons, including memory in the crop model or lurking covariates, the most likely explanation here is that the regressors used, mean growing-season temperature or precipitation, do not fully describe the conditions that affect crop yields. The mean growing-season value is only a proxy for the distribution of daily climatic conditions that crops are sensitive to, and present-day variations between years can be very different from future forced changes. Present-day variations in growing season *means* from year to year may be associated with changes in growing season *distributions* that are unrelated to changes in future warmer climates: that is, a warm year at present may be quite different from a warm year in the future [e.g. 172]. Changes in temperature distributions have been shown to strongly affect crop yields [e.g. 13, 62] though

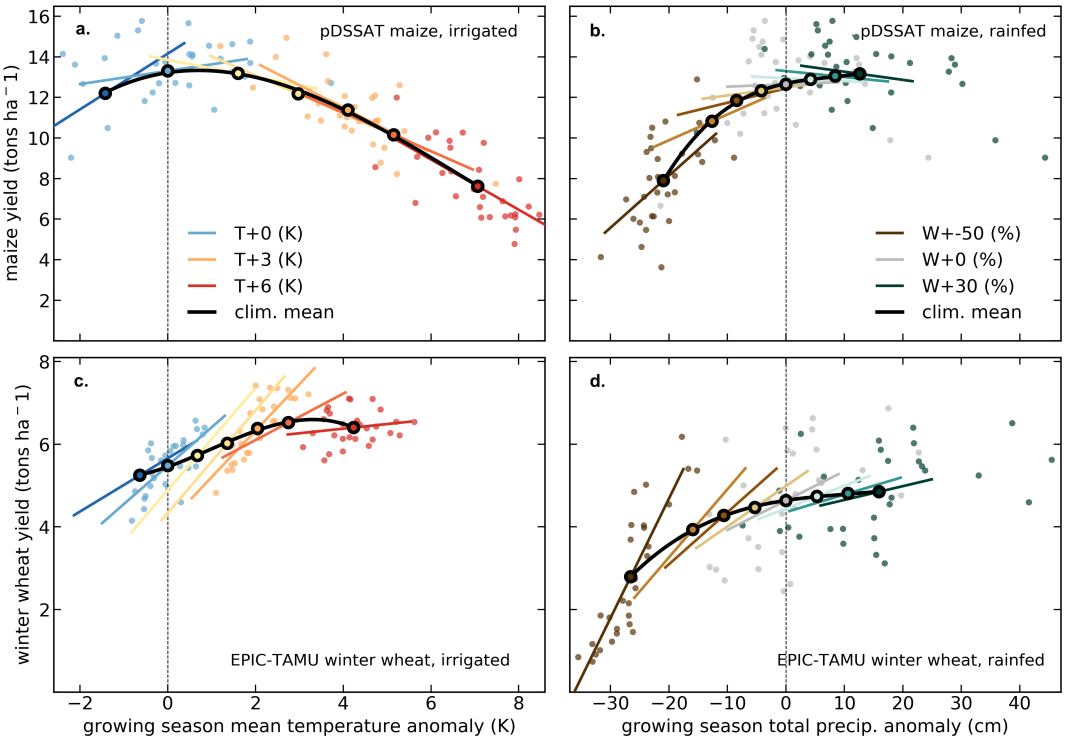


Figure 3.1: Example showing distinction between crop yield responses to year-to-year and climatological mean shifts in climate variables, showing representative high-yield regions for maize in pDSSAT (northern Iowa, top row) and winter wheat in EPIC-TAMU (France, bottom row). Left column (**a** & **c**) shows irrigated crops, all temperature cases with other variables held at baseline values, and right column (**b** & **d**) shows rainfed crops, all precipitation cases. Figure shows A0 output, in which growing seasons shift under future climate, so local growing-season temperature changes can differ from prescribed uniform offsets: for example, a 6 K applied uniform warming results in a growing season temperature warmer by ∼7 K for maize in Iowa (top right) but by less than 6 K for wheat in France (bottom right). Open black circles mark climatological mean yields and bold black lines show a third-order polynomial fit through them. Colored lines show linear regressions (by orthogonal distance regression) through the 30 annual yields of each parameter case. Colored circles show annual yields for selected cases. Differences in slopes of colored and black lines mean that responses to year-to-year fluctuations differ from those to longer-term climate shifts. Differences are generally stronger for wheat (bottom) than maize (top). Note that for rain-fed crops, slope differences in this representation could also result from correlated precipitation and temperature fluctuations in the baseline timeseries, but P-T correlations do not contribute to the effects shown here. Such correlations would complicate emulations based on year-to-year yields but would not necessarily bias them. Source: [48].

precipitation effects should be smaller since crops respond not to rainfall but to soil moisture, which integrates over weeks or even months [e.g. 148, 55, 23].

A second factor of importance is that any nonlinearity in crop responses will itself lead to

a distinction between climatological and year-to-year fits, even if distributional differences are negligible. Given the interannual variations in the climate timeseries, the mean annual yield response to a perturbation is not the same as the response of the climatological mean yield. The effect of nonlinearity may be particularly relevant for precipitation since model crop yields drop steeply and nonlinearly with increasing dryness. (Crop yields should drop under excess precipitation as well, but process-based models do not capture losses in saturated conditions well [56, 105].)

In the GGCMI Phase 2 experiment, the imposed perturbations involve no changes in underlying distributions. The choice is reasonable since climate models do not agree on distributional changes. Most models do project small mean increases in growing-season temperature variability in cultivated areas and can produce substantial local changes, but models disagree on spatial patterns. For example, in models of the Coupled Model Intercomparison Project Phase 5 (CMIP5) archive, in the the high-end RCP (Representative Concentration Pathway) 8.5 climate projections to the year 2100 [160], growing season daily maximum temperature variability over currently cultivated rice areas (weighted by production) increases by 10% in HadGEM2-ES but only by 0.4% in MIROC-ESM-CHEM. (See Supplemental Section S2.) We therefore explicitly test the assumption that distributional changes are not consequential for climatological mean yields: in Section 3.4, we confirm that an emulator trained on the GGCMI Phase 2 dataset can successfully reproduce yield changes under a full climate model projection.

Note that even though distributions of climate variables are unchanged in the GGCMI Phase 2 simulations, the spread in annual yields still becomes wider in highly impacted climate states, because of the nonlinearity of yield responses (Figure 3.2). In the GGCMI Phase 2 dataset, all crops except rice show greater year-to-year yield variance in conditions of extreme climate stress. (Rice is typically irrigated and experiences no water stress in simulations.) Increased variance has been noted in previous studies. For example, [198]

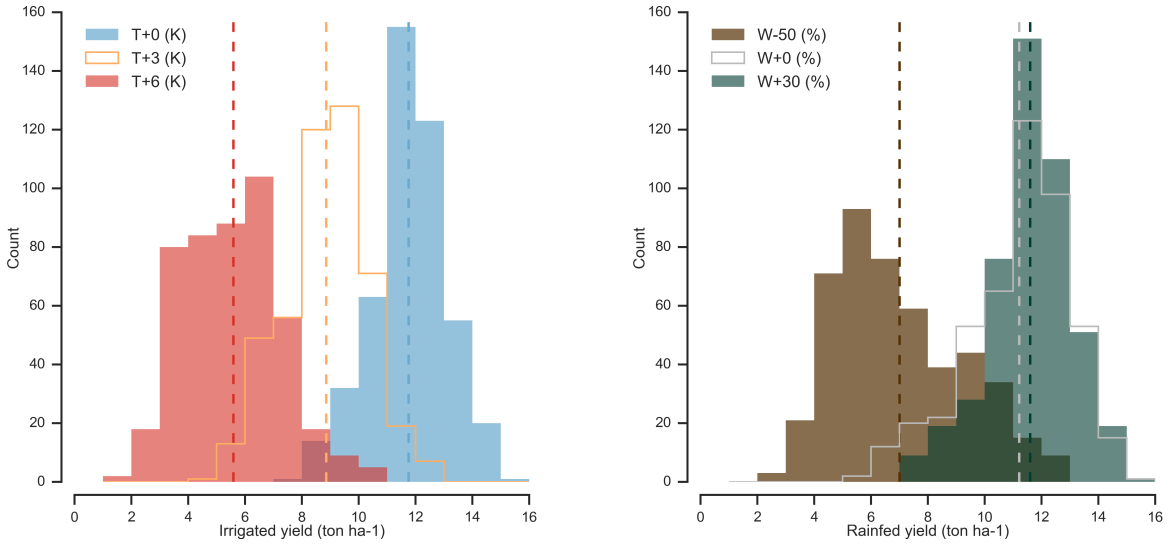


Figure 3.2: Example showing results of increased crop-yield sensitivity to year-to-year climate variations under climate stress. Yield distributions are from examples of Figure 3.1, top row, of maize in Iowa, (**left**) for irrigated maize in scenarios of altered temperature and (**right**) for rainfed maize in scenarios of altered precipitation. Because yield sensitivities rise under strong warming or drying, distributions of year-to-year crop yields widen in T+6 and P-50% scenarios relative to present-day simulations, even though all input climate timeseries have identical variance for temperature. Note: precipitation changes have different variance since the perturbations are fractional. Source: [48].

used statistical models trained on present-day yields to find a projected future increase in yield variance of U.S. maize of 20% per degree K temperature rise. Although the authors do not diagnose a specific cause of that increase, they discuss multiple potential mechanisms, including nonlinearity in responses.

3.3 Emulation

Emulation involves fitting individual regression models from GGCMI Phase 2 output for each crop and model and 0.5 degree geographic pixel; the regressors are the applied perturbations in [CO_2], temperature, water, and nitrogen (CTWN). Here, we largely discuss emulations of climatological mean crop yields with no growing season adaptation (A0 scenarios) but note that any output of the crop models can potentially be emulated. We provide separate emulations of irrigated and rainfed yields and applied irrigation water (pirrww in mm yr^{-1})

in both A0 and A1 scenarios, meaning that each model and crop combination results in 6 sets of regressions. See Supplemental Material Section S2 these additional emulation cases.

3.3.1 Statistical Model

For the statistical model of crop yields as a function of CTWN, we choose a relatively simple parametric model with a third-order polynomial basis function (Equation 3.1). If the climatological mean response is relatively smooth, then a simpler form provides a reasonable fit that allows for some interpretation of resultant parameter weights. A relatively simple parametric form also allows fast model emulation at the grid-cell level, rather than requiring spatial aggregation. Emulating at the grid-cell level preserves the spatial resolution of the parent models and means that emulators indirectly include any yield response to geographically distributed factors such as soil type, insolation, and the baseline climate.

The third-order polynomial CTWN model of Equation 3.1 contains 34 terms, since the N^3 term is omitted, as it cannot be fitted in a training set sampling only three nitrogen levels. To facilitate comparing emulators parameter by parameter, we hold this functional form across locations, crops, and models, except for several necessary distinctions: regressions for irrigated crops do not contain W terms, and regressions for models that do not sample the nitrogen levels omit the N terms. Results shown throughout the paper use this full specification, but we also show (in Section 3.3.2 below) that for all but two models, 11 terms can be dropped without significant reduction in emulator fidelity. The higher specification of the 34-term model aids primarily in regions where crops are not currently grown. Most modeling groups submitted a sufficiently large training set that the 34-term model can be fit with standard ordinary least squares (OLS), but for models with lower sampling, it must be fit with a Bayesian Ridge regression method. (See Section 3.4 for evaluation of the fidelity of emulators constructed with Equation 3.1.)

$$
\begin{aligned}
Y \; &= K_1 + K_2 C + K_3 + K_4 W + K_5 N + K_6 C^2 + K_7 CT + K_8 CW + K_9 CN \\
&+ K_{10} T^2 + K_{11} TW + K_{12} TN + K_{13} W^2 + K_{14} WN + K_{15} N^2 + K_{16} C^3 \\
&+ K_{17} C^2 T + K_{18} C^2 W + K_{19} C^2 N + K_{20} CT^2 + K_{21} CTW + K_{22} CTN \\
&+ K_{23} CW^2 + K_{24} CWN + K_{25} CN^2 + K_{26} T^3 + K_{27} T^2 W + K_{28} T^2 N \\
&+ K_{29} TW^2 K_{30} TWN + K_{31} TN^2 + K_{32} W^3 + K_{33} W^2 N + K_{34} WN^2
\end{aligned} \tag{3.1}
$$

In this study, we do not focus on comparing other functional forms or non-parametric models. In general, higher-order and interaction terms are expected to be important for representing crop yields. Higher-order terms are needed because crop yield responses to weather are well-documented to be nonlinear: e.g. [180] for T perturbations and [67] for W (precipitation). Interaction terms are needed since the yield response is expected to depend on interactions between the major inputs. For example, [115] and [193] showed that in real-world yields (with C and N fixed), the joint distribution in T and W is needed to explain observed yield variance. Other observation-based studies have shown the importance of the interaction between W and N [e.g. 7], and between N and C [138, 130]. Some prior studies have used even more complex statistical specifications in crop model emulation: for example, [15] and [14] use a 39-term fractional polynomial and "borrow information across space" by fitting grid points simultaneously across soil region in a panel regression. The GGCMI Phase 2 dataset allows fitting a simple third-order polynomial form independently at each grid cell while still providing a satisfactory emulation for all models and crops.

3.3.2 Feature Importance and Reduced Statistical Model

Because a simpler statistical model may improve the interpretability of its parameter weights, we also develop a reduced 23-term version that is satisfactory for most models and crops

(Equation 3.2, with the 11 removed terms shown in gray. To identify terms that can be omitted, we apply a feature selection cross-validation process in which terms in the polynomial are tested for importance. Higher-order and interaction terms are successively added to the regression model, and in each case we calculate an aggregate mean absolute error (weighted by currently cultivated area) and eliminate those terms that do not contribute significantly to reducing error. The procedure is illustrated in Figure 3.3. We develop a reduced statistical model by considering yields over currently cultivated land in three models: two that provided the complete set of 672 rainfed simulations, i.e. without the W_∞ simulations, (pDSSAT, EPIC-TAMU), and one that provided the smallest training set (121 input combinations, PEPIC). Although models exhibit different absolute levels of error, all three agree remarkably well on feature importance, i.e. on which terms reduce error and which provide no predictive benefit. Agreement is indicated by matching line slopes in Figure 3.3.

$$\begin{aligned} Y &= K_1 + K_2C + K_3T + K_4W + K_5N + K_6C^2 + K_7CW + K_8CN + K_9T^2 \\ &+ K_{10}TW + K_{11}TN + K_{12}W^2 + K_{13}WN + K_{14}N^2 + K_{15}CN^2 + K_{16}T^3 \\ &+ K_{17}T^2W + K_{15}CN^2 + K_{16}T^3 + K_{17}T^2W + K_{18}TW^2 + K_{19}TWN \\ &+ K_{20}TN^2 + K_{21}W^3 + K_{22}W^2N + K_{23}WN^2 \end{aligned} \tag{3.2}$$

The eliminated terms include many of those in C: the cubic; the CT, CTN, CTW, and CWN interaction terms; and all higher-order interaction terms in C. Finally, we eliminate one second-order interaction term in W and two in T. Implications of this choice include that nitrogen interactions are complex and important and that water interaction effects are more nonlinear than those in temperature. Note that some terms that did not reduce the aggregate error must still be included if a higher order version of that term provides benefit: for example, including the T^3 term requires also retaining T^2 and T terms. The

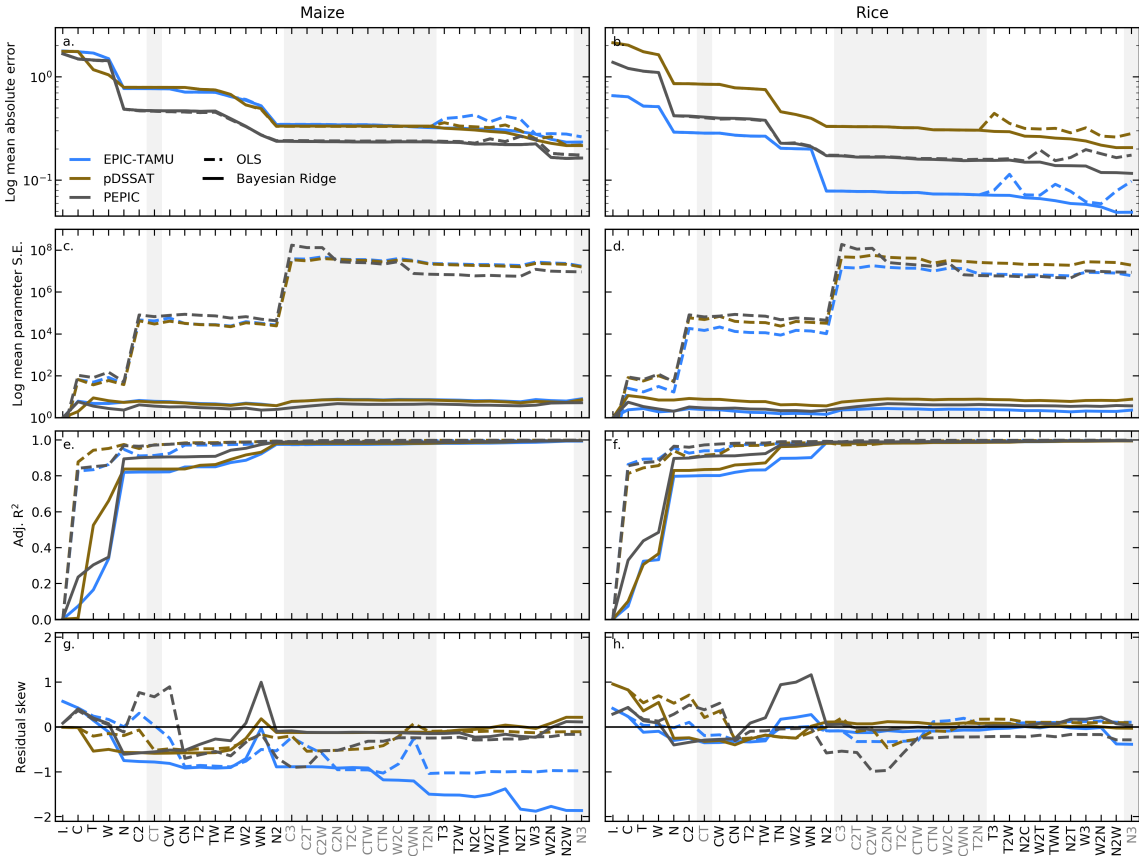


Figure 3.3: Illustration of results from the polynomial feature selection process for three different crop models (colors), for all grid cells with more than 1,000 ha cultivated for maize (**left**) and rice (**right**). Solid lines are Bayesian Ridge regression results, and dashed lines are for standard OLS. Rows show four metrics of fit quality, and x-axes show the terms successively tested in the statistical model, sequentially added to the model in order from left to right. Terms that do not reduce the aggregate error are marked in gray and are not included in the final model. **a & b:** log mean absolute error between emulated yield and simulated values calculated with a threefold cross-validation process, where the emulator is trained on two-thirds of the data and predicts the remaining third. **c & d:** log mean standard parameter error. The Bayesian Ridge method strongly reduces parameter error and results in more stable estimates. **e & f:** adjusted R^2 score for the fit at each model specification. **g & h:** distribution of the residuals. Skewness is low at the high model specifications tested in all model cases other than EPIC-TAMU maize. Source: [48].

reduced-form emulator is acceptable across currently cultivated land for all model and crop combinations other than JULES soybean and spring wheat and PROMET soybean and rice. These cases involve yield responses that benefit strongly from inclusion of higher-order carbon dioxide interaction terms. Additional terms in the statistical model also help emulation in some geographic locations outside of currently cultivated regions, where yield responses are

often non-standard. See Supplemental Material Section S2.7 for evaluation of the fidelity of emulators constructed with Equation 3.2 and for more details on JULES and PROMET.

3.3.3 Model Fitting

To fit the parameters K, we use a Bayesian Ridge regularization method [116] rather than ordinary least squares (OLS). The Bayesian Ridge method reduces volatility in parameter estimates when the sampling is sparse, by weighting parameter estimates towards zero, allowing the use of a consistent functional form across all models and locations. The choice slightly reduces mean absolute error for some of the high-order interaction terms in the model (Figure 3.3, top row) but drastically reduces standard parameter error in the model by stabilizing the estimates (Figure 3.3, third row). The estimation method scores relatively lower on adjusted R^2 for the simplest parameter specifications but quickly reaches parity with the OLS. We use adjusted R^2 as a metric because additional terms are penalized (Equation 3.3, where n is the number of samples and k is the number of features):

$$R^2_{adj} = 1 - \frac{(n-1)\cdot(1-R^2)}{n-k} \tag{3.3}$$

We use the implementation of the Bayesian Ridge estimator from the scikit-learn package in Python.

An additional diagnostic of fit quality is the distribution of residuals: normally or near-normally distributed residuals imply that errors around the fit are random and unbiased. When fitting Equation 3.1 to the GGCMI Phase 2 dataset, the distribution of the residuals depends on the number of features included in the regression, the method for estimating the parameters, and the target distribution in the training set. The residuals are only normally distributed (pvalue > 0.05 in the Shapiro–Wilk test) for a single model, PEPIC, for any specification tested here, but their skew is relatively small except in a single case, EPIC-

TAMU maize (Figure 3.3, fourth row). While including higher-order terms in the statistical model generally reduces residual skew, for EPIC-TAMU maize it increases skew instead but also reduces the error in cross-validation, which we consider more important in the context of emulation. The residual distribution suggests that projections using the EPIC-TAMU maize emulator will tend to be biased high, but in practice, the overall magnitude of these errors is below 2% of yield changes. (See Section 4.2.)

3.4 Emulator Evaluation

In this section we show illustrations of GGCMI model yield responses to climate perturbations and evaluate the ability of the emulators to reproduce them. Model emulation with the parametric method used here requires that crop-yield responses be sufficiently smooth and continuous to allow fitting with a relatively simple functional form; in Section 4.1 we show that this condition largely holds in the GGCMI Phase 2 simulations. In section 4.2 we evaluate metrics of emulator performance and show that emulation errors – discrepancies between emulation and simulation – are generally small, especially when compared to the differences across crop models or to projected yield changes. We use the term *error* because, under the "perfect model" emulation approach, we take the simulation output to be perfect ground truth. We evaluate two separate error metrics, one more loose that incorporates information about the inter-model uncertainty, and one more stringent that tests out of sample prediction error within an individual model. For both metrics, emulation error is generally small other than in limited geographic locations, usually where crops are not currently grown. Finally, in Section 4.3, we assess the emulator's ability to reproduce crop yields in a more realistic future simulation driven by a climate model projection, and find that its performance remains satisfactory. We analyze here results using the 34-term polynomial of Equation 3.1; see Supplemental Material Section S2.7 for analogous analysis of the 23-term polynomial of Equation 3.2.

3.4.1 Yield Response

Crop yields show strong spatial differentiation across geographic regions, and emulators are able to readily reproduce these patterns. Figure 3.4 shows one example of simulated and emulated yields under current climate, using maize in LPJmL. Absolute emulation errors for this model-crop combination are low – 99.8% of grid cells have errors below 0.5 tons ha^{-1} – but emulation errors as a percentage of baseline yield can be large in areas with low potential yield and no current cultivation in the real world (e.g. the Sahara, Patagonia). These regions are not currently viable for agriculture and may never become viable even under extreme climate change. Emulator spatial skill varies across models and crops, with maize being the quantitatively easiest to emulate across all models and locations.

Yield responses to the four main drivers considered here (C, T, W, and N) are also quite diverse across locations, crops, and models, but in nearly all cases the local climatological mean responses are smooth enough to permit emulation with the functional form used here. Figure 3.5 illustrates the geographic diversity of responses within a single crop and model, for rainfed maize in pDSSAT. While the[CO_2]responses (in ton ha^{-1} ppm^{-1}) are quite similar, the precipitation response is stronger in more arid locations and the nitrogen responses appear strongly location-dependent. The heterogeneity in response supports the choice of emulating at the grid-cell level. In regions with current cultivation, yields evolve smoothly across the space sampled, and the polynomial fit captures the climatological-mean response to perturbations well. Emulators do perform poorly in a few regions that involve discontinuous or irregular yield responses. Poor performance is illustrated here with PROMET maize in northern Canada, which is too cold for maize at present in PROMET (0 ton ha^{-1} yield) but shows an abrupt rise to moderate yields once temperature rises by 4 degrees. Under these conditions, the third-order polynomial cannot fit the response, and errors are high.

Crop-yield responses in all models generally follow similar functional forms at any given location though with a spread in magnitude (Figure 3.6, which shows rainfed maize in

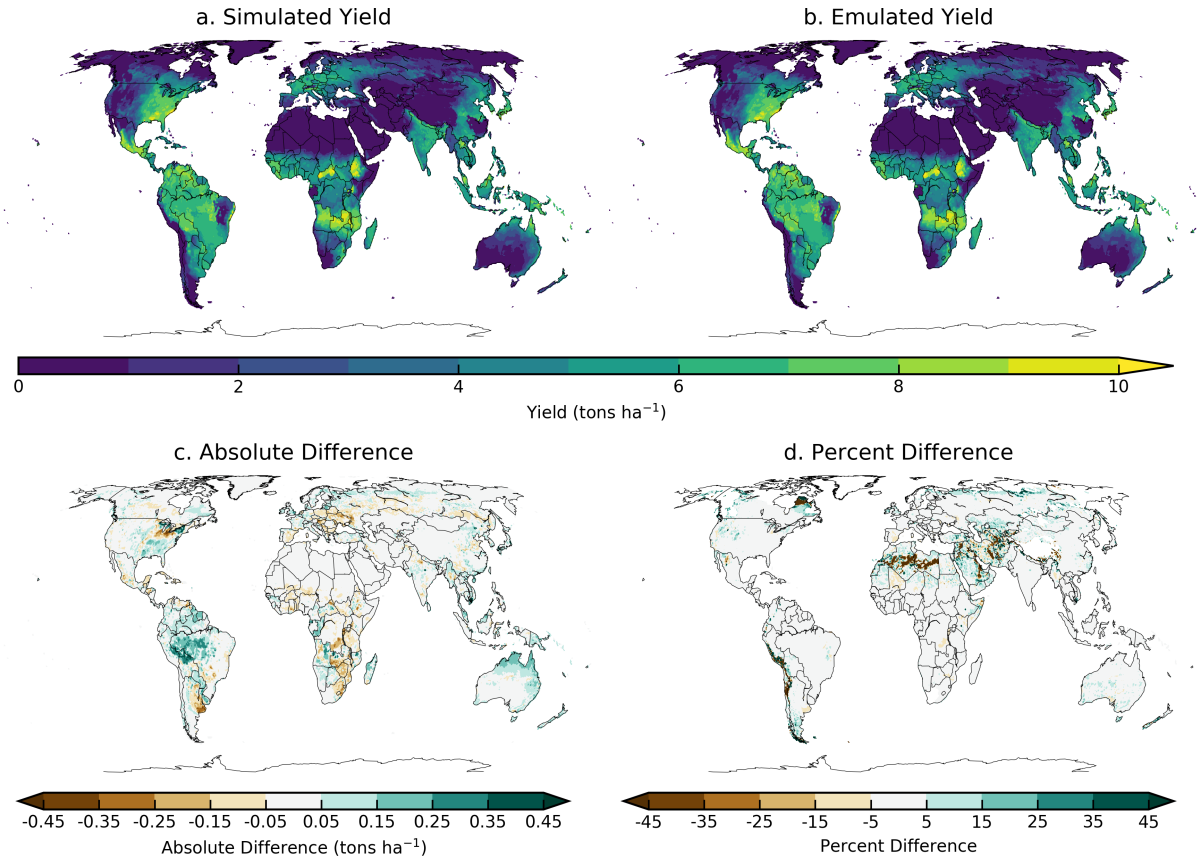


Figure 3.4: Illustration of spatial pattern in baseline yield successfully captured by the emulator: simulated (**a.**) and emulated (**b.**) yield under historical (1981-2010) conditions for rainfed maize from the LPJmL model. Absolute yield differences (**c.**) are less than 0.5 ton ha^{-1} in almost all (99.8%) grid cells across the globe. Percent difference (from simulated baseline, **d.**) is below 5% in most (75%) grid cells currently cultivated in the real world. Approximately 7% of all grid cells, but only 3% of currently cultivated grid cell, have emulated yields that differ from the baseline simulation by more than 20%. Notable exceptions include areas with very low simulated baseline yield, including for example the Sahara, the Andes, and northern Quebec. Percent error weighted by cultivation area globally is essentially zero (see also Table 3.2). Performance varies by crop and model. See Supplemental Figures in Section S2.8 for more examples. Source: [48].

northern Iowa in a selection of GGCMI models). Absolute yield differences between models can be substantial because some models are uncalibrated. In general, models are most similar in their responses to temperature perturbations and least similar to changes in [CO_2]. That is, [CO_2] fertilization effects *within* a single model are consistent across locations, but [CO_2] effects differ strongly *across* models.

Note that while the nitrogen dimension is important, it is also the most troublesome to

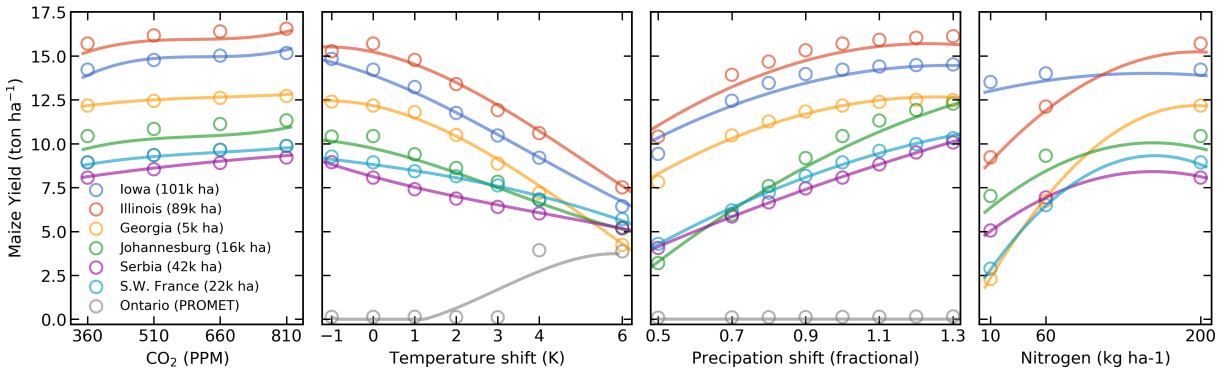


Figure 3.5: Illustration of spatial variations in yield response, which are successfully captured by the emulator. Panels show simulations (points) and emulations (lines) of rainfed maize in the pDSSAT model in six example locations selected to represent high-cultivation areas around the globe. Legend includes hectares cultivated in each selected grid cell. Each panel shows variation along a single variable, with others held at baseline values. Dots show climatological mean yields and lines the results of the full 4D emulator of Equation 3.1. In general the climatological response surface is sufficiently smooth that it can be represented within the sampled variable space by the simple polynomial used in this work. In some cases extrapolation would produce misleading results, and the emulator fails in conditions where yield response changes abruptly. Failure is illustrated here by rainfed maize in north-central Ontario for the PROMET model (in gray), which shows present-day yields of zero rising abruptly if temperature warms by 4 degrees. Source: [48].

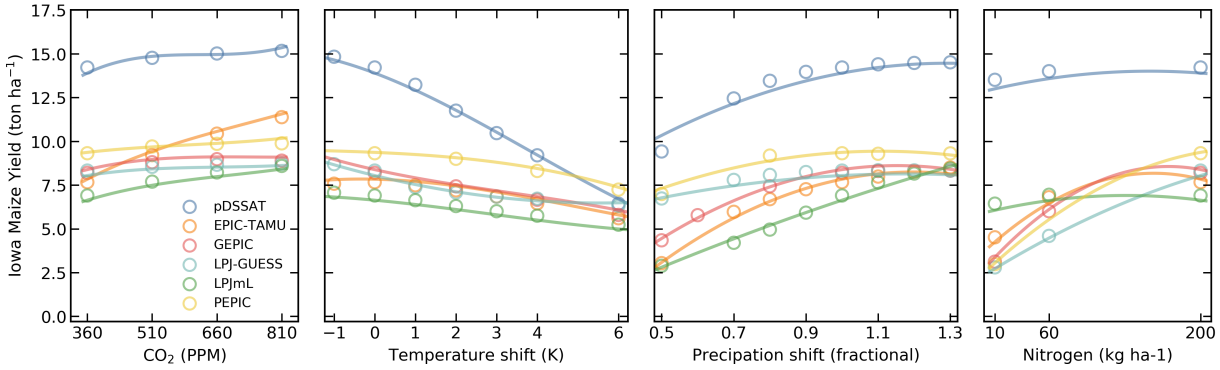


Figure 3.6: Illustration of variations in yield response across models, again successfully captured by the emulator. Panels show simulations and emulations from six representative GGCMI models for rainfed maize in the same Iowa grid cell shown in Figure 3.5, with the same plot conventions. Three models (PROMET, JULES, and CARAIB) that do not simulate the nitrogen dimension are omitted for clarity. Models are uncalibrated, producing spread in absolute yields. While most model responses can readily emulate with a simple polynomial, some response surfaces diverge slightly from the polynomial form, producing emulation error (e.g. pDSSAT here, for water), but resulting error generally remains small relative to differences across models. Source: [48].

emulate in the GGCMI Phase 2 experiment because of its limited sampling. The GGCMI Phase 2 protocol specified only three nitrogen levels (10, 60 and 200 kg N y^{-1} ha^{-1}), so a

third-order fit would be over-determined but a second-order fit can result in potentially non-physical results. Steep and nonlinear declines in yield with lower nitrogen levels mean that some regressions imply a peak in yield between the 100 and 200 kg N y^{-1} ha^{-1} levels (Figure 3.6, right). While reduced yields under high nitrogen levels are physically possible and could reflect over-application at particular times in the growing period, they are implausible at the magnitude shown here and likely an artifact of the fit. The Bayesian Ridge estimator mitigates the 'peak-decline effect' in the nitrogen dimension relative to ordinary least squares but does not entirely remove it. The polynomial fit also cannot capture the well-documented saturation effect of nitrogen application [e.g. 82] as accurately as would be possible with a non-parametric model.

3.4.2 Emulator Performance Metrics

Our emulators collectively consist of nearly 3 million individual regressions, so developing concise performance metrics poses a challenge. No general agreed-upon criteria exist for defining an acceptable crop model emulator, so we present two different metrics below, one relatively loose and one more stringent. Both metrics assess the ability of the emulator to reproduce simulated crop yields in the GGCMI Phase 2 experiment. In this section we show only results from emulators based on the 34-term Equation 3.1; see Supplemental Material Section S2.7 for analogous assessment of emulators based on the 23-term Equation 3.2.

1. Normalized error. We take as the first metric what we term the "normalized error", which compares the fidelity of an emulator to the inter-model spread. For a multi-model comparison exercise like GGCMI Phase 2, a reasonable though loose emulator criterion is that its errors be small relative to inter-model differences. The normalized error e is defined separately for each C,T,W,N scenario s as the difference between emulated and simulated fractional yield changes, normalized by the standard deviation in simulated changes across all models:

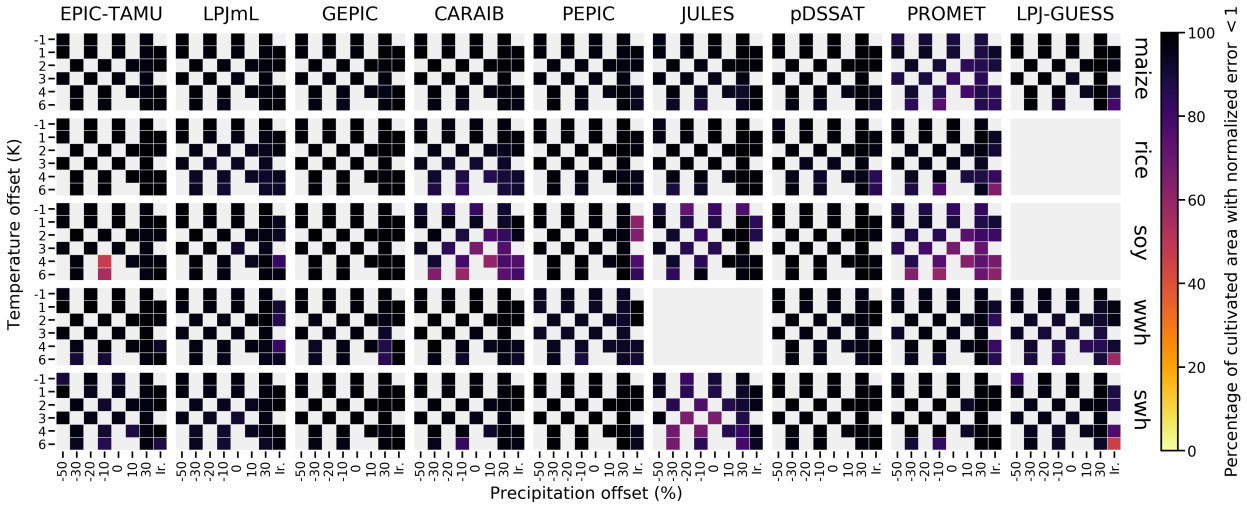


Figure 3.7: Assessment of emulator performance over currently cultivated areas based on normalized error (Equations 3.5). We show performance of all 9 models emulated, over all crops and all sampled T and W inputs ("ir." indicates the irrigated W_∞ setting) but with $[CO_2]$ and nitrogen held fixed at baseline values. Large columns are crops, and large rows are models; squares within are T, W scenario pairs. Colors denote the fraction of currently cultivated hectares ("area frac") for each crop with normalized area e less than 1 indicating the error between the emulation and simulation less than one standard deviation of the ensemble simulation spread. Of the possible 63 scenarios at a single $[CO_2]$ and N value, we consider only those for which all 9 (8 for rice, soybean, and winter wheat) models submitted data (Figure S17) so the model ensemble standard deviation can be calculated uniformly in each case. JULES did not simulate winter wheat and LPJ-GUESS did not simulate rice and soybean. Emulator performance is generally satisfactory, with some exceptions. Emulator failures (significant areas of poor performance) occur for individual crop-model combinations, with performance generally degrading for colder and wetter scenarios. Source: [48].

$$e_s = \frac{F_{em,\,s} - F_{sim,\,s}}{\sigma_{sim,\,s}} \tag{3.4}$$

where F is the fractional change in yields Y between scenario s and baseline b:

$$F_s = \frac{Y_s - Y_b}{Y_b} \tag{3.5}$$

We calculate the mean error for each grid cell, model, and crop in each C,T,W,N scenario by comparing emulated and simulated yields. A normalized error $e < 1$ means that any

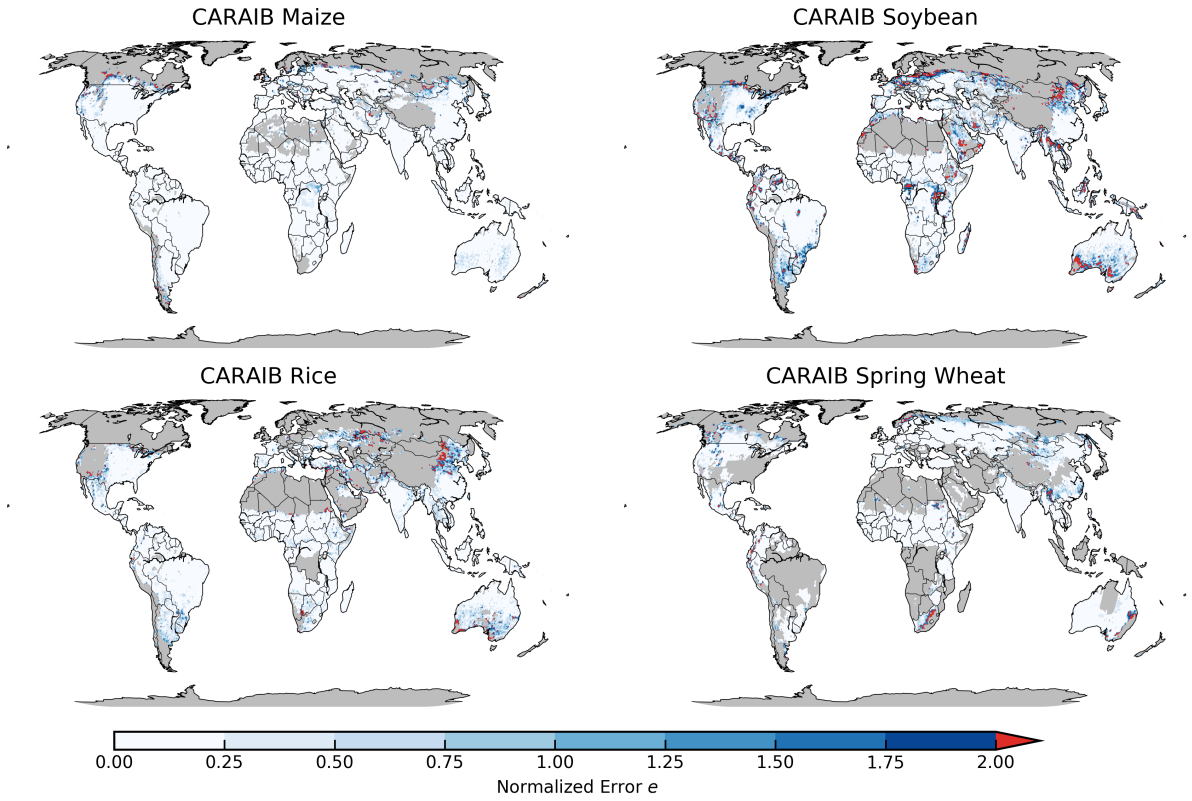


Figure 3.8: Illustration of the first test of emulator performance, applied to the CARAIB model for the T+4 scenario for rainfed crops. Colors indicate the normalized emulator error e, where $e > 1$ means that emulator error exceeds the multi-model standard deviation. For consistency, we show e only for geographic areas simulated by at least six models and where baseline yields are greater than 0.5 ton ha^{-1}. Emulator performance is generally good relative to model spread in areas where crops are currently cultivated (compare to Figure S18-S19) and in temperate zones in general; emulation issues occur primarily in marginal areas with low-yield potentials. Source: [48].

deviation of the emulation from the simulation is less than one standard deviation of the inter-model spread.

Evaluation of this metric implies that GGCMI Phase 2 emulators are generally satisfactory. Emulator performance is illustrated in Figure 3.7, which shows all models and crops crops over currently cultivated area. Over all crops and models, the average normalized error $e < 1$ over 95% of currently cultivated area. For maize, the most tractable crop to emulate, all 9 models return $e < 1$ over 97% of currently cultivated area. Only three crop-model combinations are problematic, returning $e < 1$ over less than 90% of cultivated area even when using the 34-term statistical model: PROMET and CARAIB for soybeans (79% and

83%), and JULES for spring wheat (85%). Misfits typically occur when models show strong discontinuities in yield response (as shown in Figure 3.5), or when carbon dioxide fertilization gains interact nonlinearly with changes in temperature or water. Including higher-order C terms helps in the latter case but does not reduce emulator errors to zero. See Supplemental Figures S38-S39 for examples of worst-case emulator failures.

While Figure 3.7 shows only currently cultivated land, performance can be worse in locations where crops are not currently cultivated or on marginal lands where current potential yields are low. In general, emulator performance is poor anywhere that models show steep yield changes once some threshold has been reached. Some of these cases involve complete crop failures in a changed climate, but most involve yield improvements: abrupt gains in regions that are too cold or dry under current conditions but that become viable given warming or wetting. Figure 3.8 illustrates this effect for CARAIB in the T+4 scenario, showing normalized error over all simulated areas with non-zero baseline yield and at least six models providing simulations. CARAIB emulator performance is generally good where crops are grown but can be poor ($e > 2$) in arid or mountainous zones, e.g. the edges of the Sahara, inner Mongolia, South Africa and southern Australia. Effects will vary by crop model as they differ in process implementations; see the different model description papers referenced in Table 3.1 for more details. Note that the choice of statistical model for emulation involves a trade-off in the spatial pattern of errors. The 34-term statistical model used here maximally improves emulator fidelity in problematic "fringe" areas at the expense of lowering it slightly over high-yield areas. For example, over currently cultivated land, CARAIB maize emulators have normalized error $e < 1$ over 98.5% of area with the full 34-term Equation 3.1 but over 98.8% with the reduced 23-term Equation 3.2. The effect is reversed over uncultivated land, with CARAIB maize emulators showing $e < 1$ over 93.7% of area with the full Equation 3.1 but only over 88.7% of area with the reduced Equation 3.2.

The normalized error assessment is relatively forgiving for several reasons. First, it is

an in-sample validation, with the emulation evaluated against the simulations actually used to train the emulator. Had we used a spline interpolation, the error would necessarily be zero. Second, the metric scales emulator fidelity not by the magnitude of yield changes in the evaluated model but by the spread in yield changes across models. The normalized error e for a given model then depends on the particular suite of models considered in the intercomparison exercise. The rationale for the choice is to relate the fidelity of the emulation to the true uncertainty, which we take as the multi-model spread, but the metric then has the property that where models differ more widely, the standard for emulators becomes less stringent and vice versa. In GGCMI Phase 2 the effect is manifested in the higher normalized errors for soybeans across all models, which result not because soybean yields are difficult to emulate but because models agree more closely on yield changes for soybeans than for the other crops.

2. Out-of-sample validation. We provide a second, more stringent test of emulator performance via a threefold cross validation (also termed an out-of-sample validation). In this test the GGCMI Phase 2 dataset is split randomly into two parts, with 90% of the data used to train (calibrate) the model and the held-out 10% used to test (evaluate) the fidelity of the resulting emulator. The procedure is repeated three times; in each case we calculate the root mean square error (RMSE) between the emulated (predicted) and actual simulated test set values, and then average the three results. The result is a single metric for each grid cell for each model-crop combination. As a last step, we normalize the error metric for each grid cell by dividing by its maximum yield change over the entire CTWN dataset. (Since all models have submitted the extreme T+6 scenario, this normalization choice is not problematic.) Note that this validation exercise is independent of the procedure for generating the final published emulator values, which are generated using the full CTWN dataset.

The resulting error metric is generally low. Table 3.2 shows the yield-change-normalized

RMSE for rainfed crops in all models over currently cultivated land, both in selected major producing regions and in the global average. We include all simulations in the CTWN space and report the average error value in Table 3.2. Global mean grid-cell RMSE is below 5% of maximum yield changes in all cases, or in absolute terms less than 0.2 ton ha^{-1} for all except JULES soybean simulations (0.36 ton ha^{-1}). Emulators for rainfed and irrigated crops have similar fractional errors, but since irrigated crops experience lower yield changes across the CTWN scenarios, they also have lower absolute errors. See Supplemental Material Section S2.9 for maps of cross validation RMSE for each crop and model.

Table 3.2: RMSE of emulator replication of simulated yields of rainfed crops, stated as a percentage of simulated yield change. Values are the mean grid cell error as a percentage of simulated yield change, over all currently cultivated grid cells weighted by cultivation area, for selected major regions (NA: North America, SA: South America). For comparison, global mean values are shown in parentheses. Errors are calculated using the 90-10 cross validation scheme described in text, with the model trained on 90% of the data and validated on the held-out 10% (repeated twice). All fits are made with the Bayesian Ridge method; for context, we mark with * those cases where the Bayesian Ridge is required because the OLS linear model fails (e.g. PEPIC, which has the lowest number of samples at n=121).

Model	NA Maize	SA Soybean	SE Asian Rice	NA S. Wheat	Eu. W. Wheat
CARAIB	0.7 (0.9)	2.4 (2.4)	2.4 (2.4)	1.3 (1.4)	2.7 (1.9)
EPIC-TAMU	2.4 (1.8)	1.8 (2.6)	1.6 (1.6)	1.8 (1.9)*	1.1 (1.1)
JULES	2.6 (2.6)	4.6 (4.0)	1.6 (1.7)	2.0 (2.2)	NA
GEPIC	2.1 (2.4)	1.0 (1.2)	2.0 (2.1)	3.7 (3.3)	4.0 (2.9)
LPJ-GUESS	1.0 (1.1)	NA	NA	1.0 (1.3)	1.0 (1.2)
LPJmL	1.8 (1.8)	1.1 (1.3)	1.2 (1.1)	0.8 (1.1)	1.5 (1.3)
pDSSAT	1.9 (1.7)	1.2 (1.1)	1.7 (1.6)	1.1 (1.3)	1.4 (1.5)
PROMET	3.4 (2.7)*	2.0 (2.7)*	2.1 (1.8)*	4.3 (3.7)*	4.6 (3.4)*
PEPIC	1.8 (1.8)*	1.4 (1.9)*	1.4 (1.4)*	2.3 (2.3)*	4.9 (2.9)*

Note that this relatively simple metric may be over-conservative. The randomized sampling protocol for dividing training and test sets can mean that a training set omits edge simulations at the highest or lowest value in CTWN space. The test prediction then involves extrapolating out of the training set range (e.g. predicting a T+6 case when the training set extends only to T+4), an improper use of an emulator. RMSE values would be lower if we had used a more careful sampling strategy that precluded extrapolation (e.g. "leave-

one-out"). For additional discussion of more detailed potential evaluation metrics, see e.g. [20].

3.4.3 Emulation of Realistic Climate Projections

Finally, we test the ability of an emulator based on the GGCMI Phase 2 perturbed mean training set to reproduce the response of a crop model driven by a realistic, evolving climate scenario. These emulators are trained only on growing-season means, and the GGCMI Phase 2 exercise involved only changes in means. We therefore seek to assess whether changes in the higher moments of temperature and precipitation distributions in a climate projection might have effects that lead to significant emulator error. Note that we are not asking whether year-to-year climate variability matters to crop yields; this point is well-established [156]. The question instead is whether a realistic future climate projection involves *changes* in variability large enough that they compromise an emulator based on the GGCMI Phase 2 dataset.

To assess this potential error, we generate new crop model simulations using the LPJmL crop model (taken as a representative of GGCMI models), driven by a climate simulation from the the Coupled Model Intercomparison Project Phase 5 (CMIP5) archive [92, 195, 192]. To maximize any potential bias, we choose a climate model (HadGEM2-ES) that exhibits relatively large changes in growing-season temperature variability among CMIP5 members (Supplemental Table S2), and use the high-end RCP 8.5 scenario. We also hold [CO_2] fixed to emphasize the results of temperature and precipitation changes, in the absence of the beneficial effects of increased [CO_2]. We then compare the resulting simulated yields to the output of the GGCMI LPJmL emulator driven by the HadGEM2-ES yearly growing-season T and P anomalies (Figure 3.9). The GGCMI LPJmL emulator is able to capture the yield changes well: for all crops, emulated and simulated global production in the last decade of the simulation are identical to within 1.5%. These results imply that globally, the results of

future distributional shifts on climatological yields are small relative to the effects of mean changes (Figure 3.9). The GGCMI LPJmL emulators also reproduce decadal variations in yields, which are especially strong in spring wheat grown in northern latitudes (Figure 3.9, right) and even capture much of the residual year-to-year yield variability: R^2 of emulated vs. simulated annual yield anomalies relative to the 10-year running mean is 0.8 for spring wheat (and ∼0.3 for all other crops).

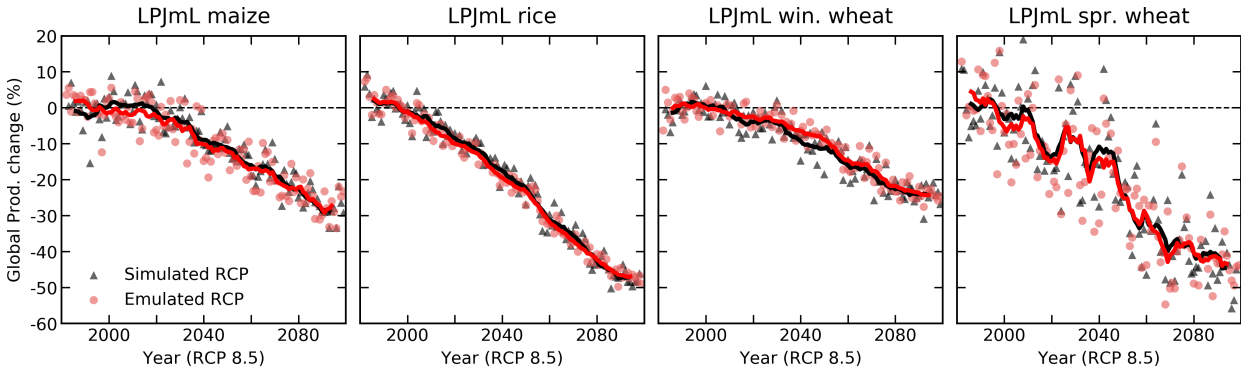


Figure 3.9: Test of emulator performance in reproducing yield simulations made with a realistic climate projection. Panels show simulated (black) and emulated (red) global production for four crops from the LPJmL model, driven with temperature and precipitation outputs from the HadGEM2-ES climate model for the RCP8.5 scenario. In both cases nitrogen and $[CO_2]$ are held fixed, at 200 kg ha^{-1} and 360 ppm. Points show yearly global production change from the 1981-2010 baseline, and lines show a 10-year running mean. See text for discussion of relating the HadGEM2-ES temperature timeseries to the appropriate offset used in emulation. Emulators trained on uniform climatological offsets reproduce well the simulated production response under a realistic climate scenario: yields at end of century match to within 1.5%. Source: [48].

Distributional effects might be expected to be stronger at high latitudes because temperature and precipitation variability are larger there, so changes in variability can be correspondingly more important. However, we find that most crops (spring wheat, winter wheat, and maize) show no emulator bias that grows with latitude. Rice is the exception: the climatological-mean emulator slightly over-predicts yield losses in the tropics and under-predicts losses at higher latitudes (where little rice is currently grown). Poleward of 30 degrees latitude, the LPJmL simulation under the HadGEM2 RCP scenario shows a 49% reduction in rice yields by end-of-century (without growing-season adaptation), but

the GGCMI-based emulator produces a reduction of only 39% (Supplemental Figure S27). These losses are concentrated in the lower mid-latitudes: only 21% of global rice is cultivated poleward of 30 degrees and only 1% poleward of 45 degrees.

It is worth noting two complications involved in comparing emulated to simulated yields under a realistic climate change scenario, as in Figure 3.9. First, it is not trivial to choose how to relate temperature or precipitation in the evolving climate scenario to the T and P offsets used as regressors to the emulator. Using growing-season mean temperature can lead to complications if crop models assume that growing season lengths shift under climate change. For consistency, we match the temperature changes in the climate scenario to their equivalent emulator regressors by calculating means over the fixed baseline growing season. This choice ensures that the emulation is appropriately matched to the simulation. Second, although the emulator outputs an estimated yield change, the baseline from which that yield change is calculated will be different between simulation and emulation because the historical climate timeseries are not identical. For example, the baseline (1981-2010) yield of winter wheat simulated by LPJmL using the AgMERRA timeseries as part of GGCMI Phase 2 is 7% lower than that simulated using the HadGEM2-ES timeseries. To minimize the effects of different historical climate assumptions, we drive the emulator with the anomaly of the climate scenario from its own 1981-2010 mean. Bias in the historical climate timeseries could in theory produce discrepancies between emulated and simulated yield changes because of the nonlinearities discussed in Section 2.2, but the effect appears to play little role in the LPJmL comparison of Figure 3.9.

3.5 Emulator Results and Products

The crop model emulators developed here can be used for a variety of applications because the emulator transforms the discrete simulation samples into a continuous response surface at any geographic scale. One use is construction of continuous agricultural damage functions

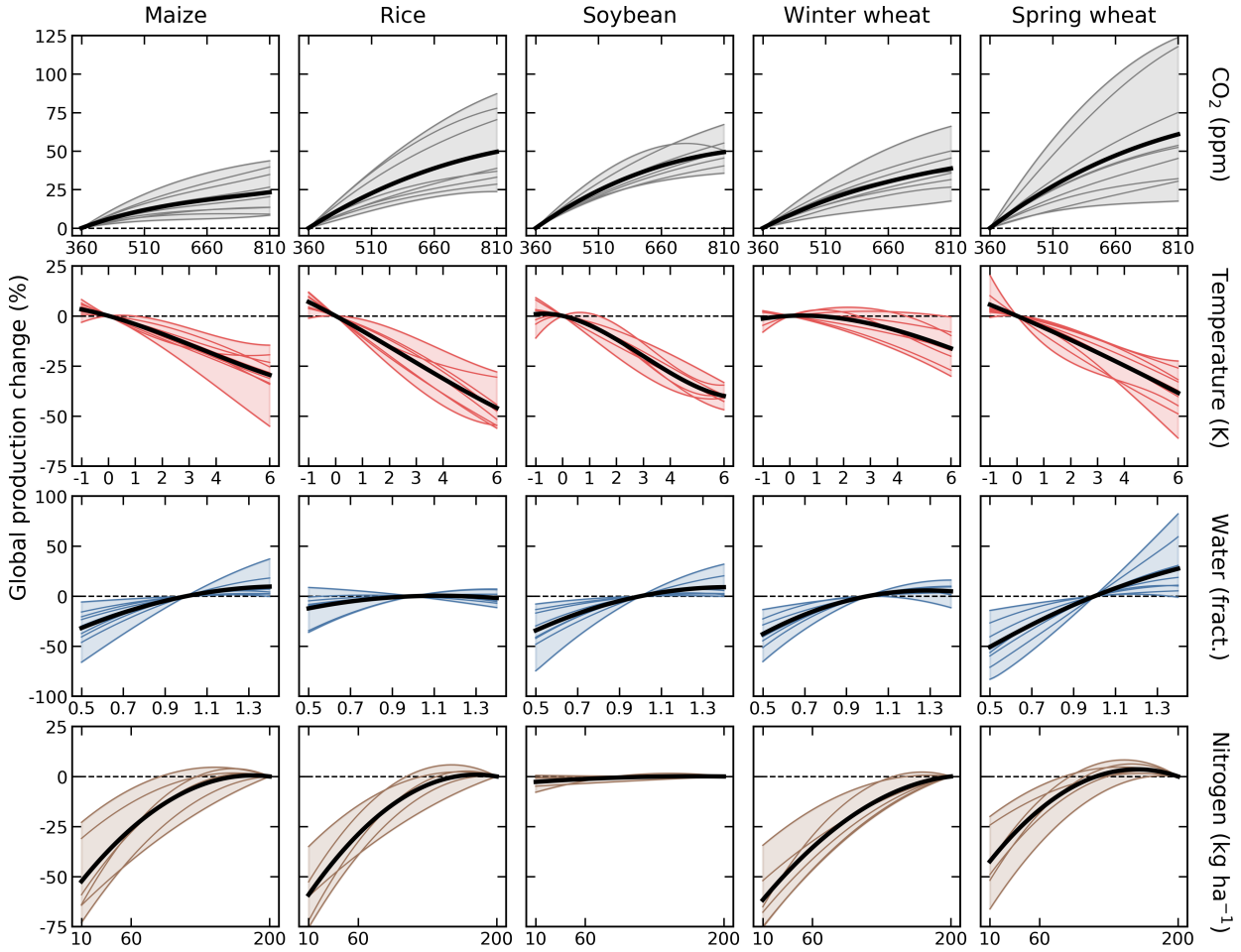


Figure 3.10: Emulated global damage functions for the five crops over the four CTWN dimensions varied in GGCMI Phase 2. Black line shows the multi-model mean and shaded area, and colored lines show the individual models. The number of models in each case varies because some models did not provide all crops or simulate the N dimension. Each panel shows response to one covariate for rainfed crops, with all others held constant at baseline values (e.g. C = 360 ppm, N = 200 kg ha^{-1}). Damages are reported as percent change in global production over currently cultivated land relative to the 1981-2010 baseline. Note that y-axis ranges are not uniform. As expected, the N response is smallest in soybeans, which are nitrogen fixers, and the C response is smallest in maize, a C4 crop. See Supplemental Figure S28 for an analogous figure identifying each crop model and Supplemental Figure S29 for damage functions for the A1 (adaptive growing season) emulators, which have reduced temperature responses. Source: [48].

in a flexible format. As an example, we present in Figure 3.10 global damage functions over each of the four dimensions tested in this study, constructed from the 4D emulation of each crop model.

These damage functions are useful in diagnosing commonalities and differences in the

responses of crop models. In most cases, models agree on the sign of responses to individual factors, but the spread in model responses is comparable to the median response. Inter-model spreads are largest for spring wheat and smallest for soybeans, as also shown in Figure 3.7. Model responses to individual factors conform to expectations. As expected, the [CO_2] response is smallest for maize, which is a C4 crop, and the nitrogen response is smallest for soybeans, which are efficient fixers of atmospheric nitrogen. Nitrogen responses in crops other than soybeans are relatively similar, and most models show saturation beginning at values less than 200 kg ha^{-1}. In nearly all crop models and for all crops except spring wheat, damages from reduced precipitation exceed benefits from increased precipitation. Spring wheat is the exception, likely because it is grown in high latitudes where rainfall may be limiting. Rice, by contrast, which is generally grown in locations with abundant water, shows nearly no benefit from increased precipitation. Note that these damage functions do not consider whether increased precipitation might permit cultivation in new areas, and also that crop models generally do represent damages from excess soil moisture well [105].

The GGCMI Phase 2 emulators are also intended as a tool for impacts assessments. The T and W functions presented in Figure 3.10 are not true global projections because they emulate the consequences of uniform shifts across the globe. However, the emulator allows building analogous damage functions based on climate model output, which has more realistic spatial patterns of changes in temperature and precipitation. In Figure 3.11, we show emulated maize responses for three crop models under the RCP8.5 scenario, using output from three climate models from the CMIP5 archive. Losses are shown as a function of mean growing-season temperature over currently cultivated land. While these damages functions aggregate over all currently cultivated land, the global coverage of GGCMI Phase 2 allows impacts modelers to develop damage functions for any desired geopolitical or geographic region larger than 0.5 degrees in latitude and longitude.

The emulated responses of Figure 3.11 allow diagnosing the factors of greatest impor-

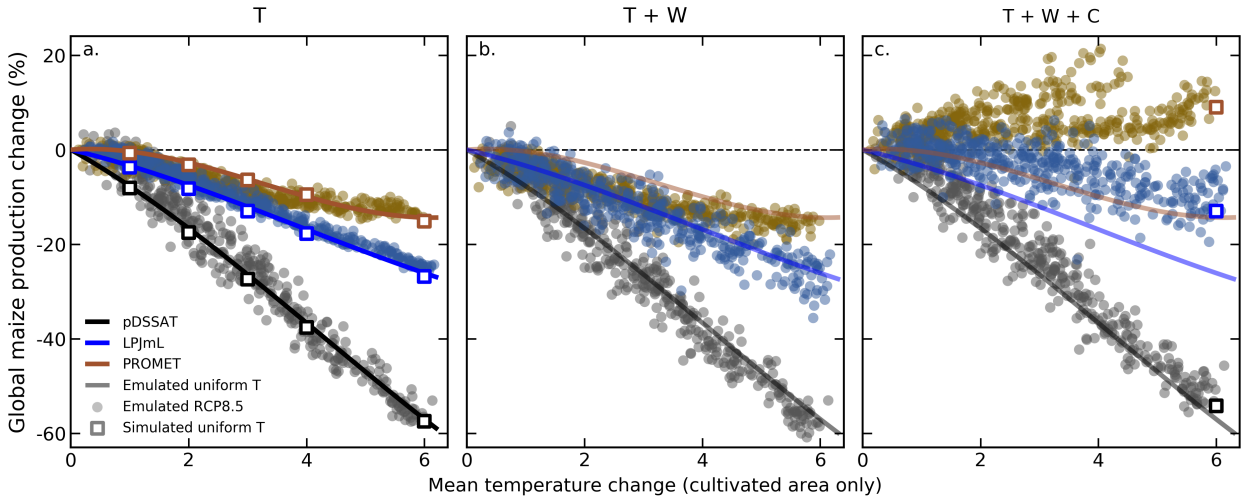


Figure 3.11: Illustration of the use of the emulator to study the factors affecting yields in a more realistic climate scenario. Figure shows emulated yield changes (relative to 1981-2010) for maize (both rainfed and irrigated) on currently cultivated land under RCP8.5 climate projections from three representative CMIP5 climate models ((HadGEM2-ES, GFDL-ESM2M, and IPSL-CM5A-LR), using changes to T only (**a**), to T and W (**b**), and to T, W, and C (**c**). The x-axis is the mean growing-season temperature change over cultivated land, computed using the historical growing season; note that these values will be higher than the corresponding global mean temperature change. Dots are emulated yearly global production changes to 2100 (90 years $\times$ 5 climate timeseries = 450 per crop model), with x-axis the mean historical growing-season T shift over all grid cells where maize is grown (unweighted by within-cell cultivated area). **a:** Using only temperature changes allows comparing regional simulated and emulated values. Open squares are GGCMI Phase 2 simulated values for each T level, with CWN held at baseline; bold lines are emulated values over uniform ΔT shifts (repeated in each panel). Emulation uncertainty (compare squares to lines) is small relative to differences across climate and crop models, and mean yield changes are similar whether T changes are applied as a uniform shift or in a more realistic spatial pattern (compare lines to dots). **b:** Adding in precipitation changes increases yield spread across climate projections and depresses yield slightly. No squares are shown in **b** because the GGCMI uniform offsets of both T and W are not directly comparable to GCM-specific changes of T and W in a climate projection. **c:** [CO_2] fertilization is small in pDSSAT, moderate in LPJmL, and very large in PROMET. The separation of groups of points in PROMET (gold) results because CMIP5 climate sensitivities differ by nearly a factor of two; points at far right are under the highest-sensitivity model, HadGEM2-ES. In RCP8.5, the 30-year-average [CO_2] at end of century is 807 ppm [159]. For comparison, open squares in **c** show GGCMI-2 simulated production changes at T+6, W=0, C=810 ppm. (Note that in these climate projections, mean [CO_2] levels when T $>$ 5.8 degrees is 912 ppm.) See Supplemental Figures S30-31 for analogous figures for other crops (spring wheat and soybeans). Source: [48].

tance to projected yield changes under future climate change. In the maize example here, temperature is the overwhelmingly dominant factor for pDSSAT, but [CO_2] responses are far larger in PROMET. [CO_2] is important across models for spring wheat, see Figure S30.

For all crop models, the aggregated effects of precipitation changes are negative, exacerbating yield losses (compare T and T+W cases) because precipitation in HadGEM2 actually declines over maize cultivation regions, especially in Central and S. America. Precipitation effects are relatively small, however, as manifested in two ways: as only a small mean shift in yield projections for individual crop models (compare T and T+W cases), and as a relatively small increase in the spread of points here at a given temperature, despite the fact that the climate projections used involve different relationships between temperature and precipitation change. By contrast, the carbon dioxide fertilization response for PROMET is so large that projections from climate models of different sensitivities ($\Delta T/\Delta[CO_2]$) become clearly separated in Figure 3.11. PROMET yield responses would be more similar if plotted as a function of $[CO_2]$ than they are when plotted as in Figure 3.11 as a function of temperature change.

Disaggregating the factors driving crop yield changes also highlights the fact that errors of emulation are much smaller than the spread across crop models or even across different climate simulations. PROMET is the most quantitatively difficult model to emulate for maize, but its comparatively large emulation error (compare open squares to lines in T case) is still smaller than the spread simply due to different T patterns across climate simulations (Figure 3.11, left, compare differences between open squares and line with the spread in dots for a given temperature value). Uncertainties in the yield damage function due to projected patterns of temperature change are in turn smaller than spread due to differing model relationships of W and T changes (Figure 3.11, middle), and for PROMET are enormously outweighed by uncertainty in climate sensitivity (Figure 3.11, right). While emulator fidelity is important to ensure, it is important to recognize that these other uncertainties will dominate any impacts assessment exercise. Note that the pattern-related yield effects are actually relatively small for maize. (In Figure 3.11, left, compare lines, which show yield changes under uniform temperature shifts, to dots, which show changes under realis-

tic warming scenarios). Pattern-related yield effects can be larger for other crops, and the uncertainties due to climate projection differences correspondingly larger: see for example soybeans in Supplemental Figure S31.

3.6 Discussion and Conclusions

In this work we describe a new class of global gridded crop model emulators for five crops (maize, soybean, rice and spring and winter wheat) and nine process-based crop models, based on the GGCMI Phase 2 dataset, a set of crop model simulations run with systematic perturbations to carbon, temperature, precipitation, and nitrogen (CTWN). The goal of this project is to provide a lightweight tool that reproduces the output of large numerical simulations of process-based crop models. The resulting emulators should provide useful tools both for diagnosing crop model behavior and for climate impacts assessment, at least of large-scale time-averaged responses.

Specific findings of this work include that:

- In crop models, the climatological mean yield responses to uniform perturbations in growing-season mean temperature and precipitation are very distinct from responses to historical weather fluctuations associated with the same mean differences. This result suggests that when emulating crop models, care must be taken if considering responses on both short and long timescales. The large GGCMI Phase 2 experiment allows us to emulate climatological-mean responses with a simple statistical model without relying on the "natural experiment" of year-to-year variations.

- Climatological mean responses in all models can be well-fit with a simple third-order polynomial in mean growing-season C, T, W, and N. The large GGCMI training set allows fitting in most cases with OLS, but use of a Bayesian Ridge regression provides additional stability and prevents overfitting. For most crop models, emulation is also

possible with a simplified version of the statistical model with only 23 terms.

- The resulting emulators are highly flexible: they capture the strong geographic difference in crop yields and yields responses, can perform well on models with quite different sensitivities to climate or [CO_2] changes. Emulators can faithfully reproduce the output of process-based crop models in both in- and out-of-sample tests. Emulation error is generally small other than in localized regions where crops are not currently grown: across all models and scenarios, errors over currently cultivated land never exceed 5% of yield changes at either global or regional scale.

- Emulators trained on the GGCMI Phase 2 dataset, which samples over uniform climate perturbations, can effectively reproduce the behavior of crop models driven by realistic future projections of future T and P changes. This result suggests that any projected changes in weather distributions (temperature and precipitation variability) have relatively little effect on crop-model yield responses relative to changes in means, at least on the regionally aggregated level.

- The GGCMI emulators should provide powerful tools for both model comparison and impacts assessments. The emulators can be used to develop standalone damage functions at any geographic scale larger than 0.5 degrees or can be integrated directly into a larger integrated assessment model (IAM) framework. Emulators can also be used to study differences across crop models in responses to individual drivers of yield changes, making them useful for model comparison and improvement.

While an emulator that captures the response of a process-based crop model in a lightweight form will never be more *accurate* then its parent model, it can have multiple advantages over a numerical simulation. Emulation over the systematic sampling of the GGCMI Phase 2 experiment provides information on the influence of multiple interacting factors in a way that individual, more realistic process-based model runs cannot. Because we use a parametric

statistical model, fitted parameter values can be physically interpreted to help understand differences between crop models. The flexibility and low computational requirements of emulators also make them particularly suitable for applications in integrated climate change impact assessments and projections of land-use change [e.g. 132]. Data storage requirements are reduced by three orders of magnitude: the yield output for a single crop model simulating all GGCMI Phase 2 scenarios for five crops is ∼12.5 GB, while the equivalent global gridded emulator parameters are only ∼20 MB. Computational requirements are nearly negligible: a thousand years of global 0.5 degree yields, i.e ∼40,000,000 individual yield projections, can be emulated in 20 seconds on a laptop computer. The resulting suite of emulators should find considerable use in climate impacts analyses [e.g. 189] and allow explicit evaluation of the uncertainty embedded in the choice of climate and crop models [126].

Several cautions should be noted when using the emulators presented here. First, extrapolation outside the GGCMI Phase 2 sample space should be avoided. Polynomial fits, while faithful within sample, quickly become non-physical outside of the tested range. This constraint is important given the strong warming expected under high-end greenhouse gas concentration scenarios (e.g. RCP8.5): if growing seasons are held fixed, climate model project mean temperature increases above 6K by end of century in many agricultural regions. Second, while the emulators are valuable for understanding the shape of yield responses and the factors that drive them, the absolute values of emulated yields should be treated with caution. The GGCMI Phase 2 models are not formally calibrated, so the emulators should be used for absolute projections only in combination with historical data. Third, neither growing season specification tested in GGCMI Phase 2 (A0 and A1) accounts for a major potential adaptation pathway under climate change, a shift to earlier or later planting dates [204], or generally different growing seasons [120]. And finally, the emulator should not be used to predict individual yearly yields, as the forced climatological mean yield response will not match the response to mean growing-season weather in a single year. The emulator

cannot provide a measure of changing yield variance and should not be used to evaluate extremes.

In summary, the GGCMI Phase 2 dataset and emulators invite a broad range of potential future avenues of analysis. Future studies using the emulators described here could include a detailed examination of interaction terms, robust quantification of model sensitivities to input drivers, and evaluation of geographic shifts in optimal growing regions. The large suite of crop models emulated lends itself particularly well to model comparison efforts, including identifying locations of model consensus (or lack thereof) and causes of model differences. While studies of yield responses to changes in growing-season variability would require new simulations, the emulators presented here provide a ready means of testing the null hypothesis that such effects are small. (Structured training sets could be constructed to directly study responses to variability changes: see e.g. [144, 65] for methods of constructing synthetic climate timeseries with altered variability.) The GGCMI Phase 2 dataset can be used as a testbed for examining the ability of statistical models that use more detailed within-season regressors to capture both year-over-year and climatological changes, and for more systematic studies of emulation itself, including evaluation of alternate statistical specifications or machine learning methods. In general, the GGCMI Phase 2 experiment demonstrates the promise and utility of systematic parameter sweeps for improving understanding of the factors driving crop responses and for evaluating and improving process-based crop models.

3.7 Data Availability

The polynomial parameters for crop model emulators are available at *https://doi.org/10.5281/zenodo.3592453.*

CHAPTER 4

AGRICULTURAL BREADBASKETS SHIFT POLEWARD GIVEN ADAPTIVE FARMER BEHAVIOR UNDER CLIMATE CHANGE

Published in *Global Change Biology* in collaboration with:

Christoph Müller[1], Sara Minoli[3], Joshua Elliott[2], Christian Folberth[4], Charles Gardner[5], Tobias Hank[6], R. Cesar Izaurralde[7], Jonas Jägermeyr[8,3,9], Curtis D. Jones[7], Wenfeng Liu[10], Stefan Olin[11], Thomas A.M. Pugh[11,12,13], Alex C. Ruane[8], Haynes Stephens[1,2], Florian Zabel[6] and Elisabeth J. Moyer[1,2]

[1]Department of the Geophysical Sciences, University of Chicago, Chicago, USA [2]Center for Robust Decision-making on Climate and Energy Policy (RDCEP), University of Chicago, Chicago, USA [3]Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, Potsdam, Germany [4]Ecosystem Services and Management Program, International Institute for Applied Systems Analysis, Laxenburg, Austria [5]Program on Global Environment, University of Chicago, Chicago, USA [6]Ludwig-Maximilians-Universität München (LMU), Munich, Germany [7]Department of Geographical Sciences, University of Maryland, College Park, MD, USA [8]NASA Goddard Institute for Space Studies, New York, USA [9]Columbia University Center for Climate Systems Research, New York, USA [10]College of Water Resources and Civil Engineering, China Agricultural University, Beijing 100083, China [11]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden [12]School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK [13]Birmingham Institute of Forest Research, University of Birmingham, Birmingham, UK

Modern food production is spatially concentrated in global "breadbaskets". A major unresolved question is whether these peak production regions will shift poleward as the climate warms, allowing some recovery of potential climate-related losses. While agricultural impacts studies to date have focused on currently cultivated land, the Global Gridded Crop Model Intercomparison Project (GGCMI) Phase 2 experiment allows us to assess changes in both yields and the location of peak productivity regions under warming. We examine crop responses under projected end-of-century warming using 7 process-based models sim-

ulating 5 major crops (maize, rice, soybeans, and spring and winter wheat) with a variety of adaptation strategies. We find that in no-adaptation cases, when planting date and cultivar choices are held fixed, regions of peak production remain stationary and yield losses can be severe, since growing seasons contract strongly with warming. When adaptations in management practices are allowed (cultivars that retain growing season length under warming and modified planting dates), peak productivity zones shift poleward and yield losses are largely recovered. While most growing-zone shifts are ultimately limited by geography, breadbaskets studied here move poleward over 600 km on average by end of the century under RCP8.5. These results suggest that agricultural impacts assessments can be strongly biased if restricted in spatial area or in the scope of adaptive behavior considered. Accurate evaluation of food security under climate change requires global modeling and careful treatment of adaptation strategies.

4.1 Introduction

While agriculture is widespread on the Earth's surface, food production is spatially concentrated: 80% of global cereals (and soybeans) and nearly half of total global calories are produced from just ∼3% of ice-free land surface area [SPAM, 84], largely concentrated in temperate "breadbaskets" or "rice bowls". From the beginnings of modern climate science, there has been concerns that climate change would drive these regions of optimal cultivation poleward, causing significant economic disruptions. Newman et al. [133] estimated that a 1°C increase in daily temperature could shift the North American maize belt poleward by 175 km, and other studies in the 1980s found similar effects [e.g. 164, 206]. A large body of work since then has continued this approach of evaluating potential changes in agricultural suitability [e.g. 1, and references therein]. While many studies use suitability rules based only on mean temperature and precipitation [e.g. 21, 97, 60], others incorporate soil characteristics [e.g. 217] or employ more complex climate suitability approaches [e.g. 66, 218, 154, 100, 71]. In economics, rule-based suitability maps have been used to investigate how future climate-driven changes in production across countries may impact the global food trade [e.g. 26, 70]. However, while climate change is already modifying the spatial distributions of many species of plants and animals [e.g. 142, and references therein], evidence for real-world shifts in agricultural growing zones is more limited [184].

Despite the concern about shifting "breadbaskets", nearly all modeling assessments of the impacts of climate change on crop yields consider production only on currently cultivated land [e.g. 129]. Many studies of future yields use statistical models trained on historical records of crop yields and weather, an approach that can only produce predictions over regions where those records exist [e.g. 180, 124, 2, 54, 101, 151]. These empirical approaches cannot test whether optimal production zones may shift into regions where crops are not currently grown. By contrast, process-based simulation models often produce global output, but in reporting production changes, typically focus on current cultivation where models are

more skillful [e.g. 165, 4, 6, 88, 128]. Both "static" approaches – empirically-based statistical models and process-based models run over currently cultivated land – tend to project significant yield losses under expected climate change when effects of [CO_2] fertilization are neglected [e.g. 220]. In the more realistic case of including [CO_2] [197], yield responses are mixed, often with different directions of change for specific regions and crops [165, 88].

In reality, the climate-induced shift in optimal production zones will likely fall between the "perfect" shift implied by the suitability approach and the stasis assumed in most impact assessments. Climate-suitability approaches still have difficulty including limiting factors such as various soil characteristics, precipitation variability, and total radiation. Global process-based models incorporate all of these physical factors, and in theory could accurately map evolving spatial patterns of crop yields, but in practice are limited by difficulties in accounting for human factors instead. Management choices such as planting dates and cultivar characteristics (which determine harvest dates) are typically specified for each geographic region based on current data and hard-wired into the model parameters. (Some previous studies have partially implemented a dynamic planting date within a fixed historical window [165].) The models therefore preclude the dynamic behavioral responses needed to capitalize on changing suitability. Typically, growing seasons are shorter and start later in the year in higher latitude production zones, reflecting the seasonal cycle. For example, maize planting starts in May in North Dakota, 1 month later than in Missouri [200], to avoid late frosts. Crop models cannot realistically capture changes in future spatial patterns of yields as the climate warms without relaxing that constraint on growing season start and length. For this reason, it is unsurprising that several recent studies using yield projections from process-based models without adaptations find little geographic shift of production under warming [189]. Even when a crop models with dynamical planting date in a fixed window is employed within an economic model [e.g 91], little production shift was found under warming.

Recent experiments with global crop model ensembles that explicitly consider growing

season adaptations now allow evaluating the factors that control poleward agricultural shifts. In this study we utilize seven process-based crop models that participated in the Global Gridded Crop Model Intercomparison Project (GGCMI) Phase 2 [34, 126, 47] to test how the shift of peak agricultural productivity regions depends on treatment of farmer adaptation.

4.2 Materials and Methods

4.2.1 Crop Model Simulations

This analysis utilizes simulations compiled as part of GGCMI Phase 2 [47, 48], an exercise involving widely-used mechanistic process-based crop models conducted within the Agricultural Model Intercomparison and Improvement Project (AgMIP) [163, 174]. Of the 12 models participating in GGCMI Phase 2, we use 7 here (Table 4.1), excluding 3 that provided insufficient simulations and 2 that used non-harmonized treatment of growing season length (see below). We also analyze an additional set of simulations by one of the GGCMI models, LPJmL (the Lund-Potsdam-Jena managed Land model), that includes more detailed growing period adaptation representation (Table 4.2). All models represent a suite of biophysical processes including photosynthetic light utilization, soil water and nutrient dynamics, phenological development, heat and water stress, evapotranspiration, and [CO_2] effects. All models simulated maize, rice, soybean, and spring and winter wheat other than LPJ-GUESS, which omitted rice and soybean. Simulations were run at 0.5 degree spatial resolution and cover over 80% of the Earth's land surface, including all cultivated land, 58% of land surface area, and much uncultivated land, excluding areas judged unsuitable for cultivation under any climate scenario. See [47] for details.

In GGCMI Phase 2, runs were conducted with uniform perturbations applied to a 30-year timeseries of historical climate inputs. Most models use daily data from the NASA AgMERRA 0.5 degree gridded reanalysis product [171], but PROMET uses sub-daily data

Table 4.1: The 7 process-based globally gridded crop models used in the study, which fall into two categories. *Site-based* models were originally developed for field-level simulations and typically involve more detail on cropping systems. *Ecosystem* or land surface models were originally developed to represent the natural terrestrial biomes and typically involve a more detailed representation of the water cycle. PROMET was originally developed from an hydrological model but is now an intermediate case.

Model	**Key citations**	**Model type**
pDSSAT	[36, 94]	Site-based
EPIC-TAMU	[86]	Site-based
GEPIC	[109, 42]	Site-based
PEPIC	[110, 111]	Site-based
LPJmL	[203]	Ecosystem
LPJ-GUESS	[107, 135]	Ecosystem
PROMET	[59, 117, 215]	Site-based/Ecosystem

from the WFDEI climate dataset [208]. The experimental protocol involved 9 levels for precipitation, 7 for temperature, 4 for [CO_2] and 3 for applied nitrogen, for a total of 672 simulations for rain-fed agriculture and an additional 84 for irrigated crops. Perturbations are uniform in space and therefor represents a possible climate future that could occur in each individual gridcell, but not one which would occur simultaneously across all gridcells. (See [48] for more details.) All simulations are run twice, first with present-day cultivar genetics and then with assumed cultivar adaptations. The additional LPJmL simulations involve more extensive treatment of growing period adaptation. These simulations are driven by climate projections from 4 CMIP5 models under the RCP8.5 scenario. Representation of growing season adaptation in both datasets is described in more detail below and in Supplementary Material Section 3.1.

In the GGCMI Phase 2 historical base case, models are harmonized to match the same region- and crop-specific planting and harvest dates [34]. For all models, crop phenology parameters (e.g. phenological heat unit requirements) were calibrated to match harmonized calendars, largely adapted from the SAGE (Center for Sustainability and the Global Environment, University of Wisconsin) crop calendar [177]. See [47] for crop calendar details

Table 4.2: The process-based crop model scenarios presented in this analysis.

Scenario	**Climate input**	**Model configuration**	**Models included**
GGCMI Phase 2 baseline **(A0)**	Uniform perturbations in temperature applied to historical reanalysis.	Phenology parameters held fixed. No adaptation measures.	pDSSAT, EPIC-TAMU, GEPIC, PEPIC, LPJmL, LPJ-GUESS, PROMET
GGCMI Phase 2 adaptation **(A1)**	Uniform perturbations in temperature applied to historical reanalysis.	Phenology parameters adjusted to maintain historical growing season length under warming.	pDSSAT, EPIC-TAMU, GEPIC, PEPIC, LPJmL, LPJ-GUESS, PROMET
No adapation	CMIP5, 4 climate models under RCP8.5.	Phenology parameters held fixed. No adapation measures.	LPJmL only
Planting advance	—"—	Planting day dynamically advanced under warming.	—"—
Harvest delay	—"—	Phenology parameters adjusted to maintain historical growing season length.	—"—
Full adaptation	—"—	Planting day dynamically advanced under warming and phenology parameters adjusted to maintain historical growing season length.	—"—
Full adaptation + uniform soils and radiation	—"—	—"— + Central Iowa soils and radiation are applied everywhere.	—"—

and Supplementary Section S1 and Table S1 for crop-specific parameters and details on the process for tuning each model. In the no-adaptation GGCMI Phase 2 simulations (termed "A0"), with phenological parameters left unchanged from the historical case, growing periods shrink under warming because crop development accelerates. In the adaptive cultivar choice simulations (termed "A1"), models are re-parameterized for each temperature level to delay crop maturity and maintain the original growing period length (averaged across each 30 year simulation). Note that the A1 runs capture only partial adaptation, since the planting date is left fixed at its historical values regardless of warming. Even so, adaptation substantially

reduces warming-related yield reductions on currently cultivated land [121, 216, this issue].

Additional LPJmL simulations involve more extensive treatment of cultivar adaptation beyond delaying maturity. The LPJmL model uses a rule-based parameter adjustment [204, 120] that can dynamically adjust both crop phenological parameters and planting and harvest dates as climate changes. Simulations shown here are run over two 20-year periods, 1985–2005 ("present") and 2080–2099 ("future"), with parameters set according to the 20-year mean climate. We explore 5 cases with the LPJmL model: 1) a *no-adaptation* case that utilizes the same historical growing seasons as the GGCMI Phase 2 protocol and keeps phenological parameters fixed, as in A0, so that growing seasons shorten under warming; 2) a *maturity delay* case in which the required accumulated phenological heat units (PHU) at maturity are adjusted to maintain growing season length, as in A1; 3) a *planting advance* case in which phenological parameters are fixed but planting date is adjusted (typically advanced) under warming according to rules based on mean temperature and precipitation [204]; 4) a *plant & maturity* case that combines the planting date change of 3) and cultivar modification of 2) [120]; and finally, for exploring the factors controlling breadbasket locations, 5) a *soils & radiation* case that combines the full *plant & maturity* adaptations of 4) with global soils and radiation set to the values in central Iowa, the heart of the North American maize breadbasket. (Soil input data are sourced from the Harmonized World Soil Database [HWSD 40], from the ISRIC-WISE database [12], or from [181] depending on individual crop model requirements, see [47] for details.) LPJmL planting date rules are location-specific and consider both temperature and precipitation. In the extra-tropics, where the dominant seasonal change is in temperature, crops are sown when the 30-year average of daily mean temperatures cross a crop-specific threshold, for example 14°C for maize and 18°C for rice. for rice (see [204] for details). In the tropics, where precipitation cycles play a larger role, crops are sown at the onset of the rainy season. See [204] for details on planting dates and [120] for details on maturity dates, though note that LPJmL was run with slightly modified

rules on maturity dates since this description. These adaptation cases are more extensive than most prior studies, but are still necessarily simplistic [e.g 24].

All LPJmL simulations are run under both dynamic [CO_2], matching that in climate projections, and fixed [CO_2]. This choice does not affect growing season timing, since the LPJmL growing-season rules are functions only of temperature and precipitation. Similarly, the GGCMI phase 2 protocol varies temperature and [CO_2] independently, and so provides simulations at a given degree of warming at a variety of [CO_2] levels from low (360 ppm) to high (810 ppm). While elevated [CO_2] can strongly affect yields, its effects are largely uniform in space, and therefore produce little change in the *relative* latitudinal distribution of potential yields. Models also disagree considerably in the impacts of [CO_2], so that varying [CO_2] in the analysis complicates the model comparison. For these reasons we predominantly show here results using fixed [CO_2], but show in Supplemental Material figures analogous figures using dynamic or elevated [CO_2]. The LPJmL fixed-[CO_2] simulations use the 2005 value of 380 ppm; in the GGCMI Phase 2 simulations, the "low" case uses the mean over the 1980–2010 period, 360 ppm. In Supplemental Material we show LPJmL simulations under the RCP8.5 scenario, where [CO_2] in the 2080-2099 "future" period ranges from 760 to 935 ppm, with a mean of 845 ppm, and the GGCMI Phase 2 case where warming of T+6°C is paired with 810 ppm [CO_2]. For comparison, we also show lower levels of warming (+2°C) at fixed [CO_2] (360 ppm) in order to discuss the robustness of the findings on the shift of peak productivity regions against these choices.

In this analysis, we use only crop model projections for the case of high nitrogen input, at 200 kgN ha^{-1}, so that current spatial patterns in fertilizer inputs do not confound the analysis. This choice implicitly assumes that farmers will be able to provide all necessary inputs to exploit the climatic potentials at any point in time.

4.2.2 Breadbasket Analysis

Breadbaskets have been defined in many ways in previous studies [e.g 8, 54]. Broadly, a breadbasket is defined as a localized area with especially high production of an individual crop within a region. However, the scale of analysis is a subjective choice, and while some methodologies result in six global "breadbaskets" for maize – North America, North China Plain, Western Europe, Northern India, Brazil, and Argentina [54] – many smaller regional areas of concentrated production are also often characterized as breadbaskets (e.g. the San Joaquin valley in California). In this analysis, we define a breadbasket as those gridcells in the top 10% of simulated yields for a given crop in a continent-scale region. This definition is in line with the scale of analysis in previous studies of North America and Western Europe [e.g. 101] though slightly larger than that in studies of South America and East Asia [e.g. 54]. The breadbasket center is taken as the mean latitude and longitude of those gridcells in the top yield decile. Analysis here focuses on a few areas of strong importance to global food supply: primarily the North American maize belt, but with comparisons to wheat in Europe, soybeans in South America, and rice in East Asia. When comparing model output to maps of real-world data, we define the real-world breadbasket as the most intensely cultivated regions equivalent to a similar total area: for the N. American maize belt, ∼165 Mha (million hectares) total area, of which 30 Mha is cultivated in maize.

For consistency, when evaluating changes in breadbasket locations under warming scenarios, we compare model-specific results with "future" and "present" simulations for each model, so that any model errors in spatial locations do not overly distort the analysis. Global gridded crop models do differ in their patterns of potential yields under present-day conditions, and in some cases appear to deviate from observed real-world cultivation patterns. For North American maize, for example, all models except pDSSAT place many of the best-yielding gridcells further south than the real-world maize belt. This mis-location of optimal growing locations appears amplified in ecosystem-type models (Supplemental Figure S50).

Model skill at simulating present-day spatial yield patterns has been less well-tested than their skill at capturing year-to-year variation in country-level production, but differences are expected due to lack of detailed input data on management practices and cultivars [e.g. 34, 126, 79, 41]. However, all models do reproduce strong spatial differences in yield and concentrated areas of optimal production.

Breadbasket temperature is reported as the mean growing period temperature in those gridcells in the top yield decile. When showing climatological future breadbasket temperature increases, we use the location of the future breadbasket, but compute temperatures over fixed historical growing seasons for consistency. Dynamic growing seasons and adaptive changes in planting date would make realized temperature increases smaller; see Supplemental Figure S73. We do not show temperature increases for wheat, because adaptations in wheat production may also involve changes from winter to spring wheat and vice versa.

4.3 Results

4.3.1 Expected Future Growing-Season Climate Changes

In a simple climate suitability framework, expected end-of-century climate change would imply substantial geographic shift of breadbaskets. To illustrate this, we show the case for maize in North America in Figure 4.1. While maize is the second-most widely grown crop globally, with almost 200 Mha [46] under cultivation, its production is highly concentrated, and the relatively small U.S. corn belt accounts for a full third of global production [205]. The center of the U.S. corn belt is the state of Iowa, which produces 2.5 $\times$ as much corn as all of Mexico [46, 200]). The compact size of the corn belt means that the vast majority of U.S. maize is grown at temperatures similar to those in Iowa: the standard deviation of corn belt mean growing season temperatures is only 2.3°C around the Iowa mean of $\sim$21°C [AgMERRA, 171] (Figure 4.1, left). In Asia, maize production on the North China

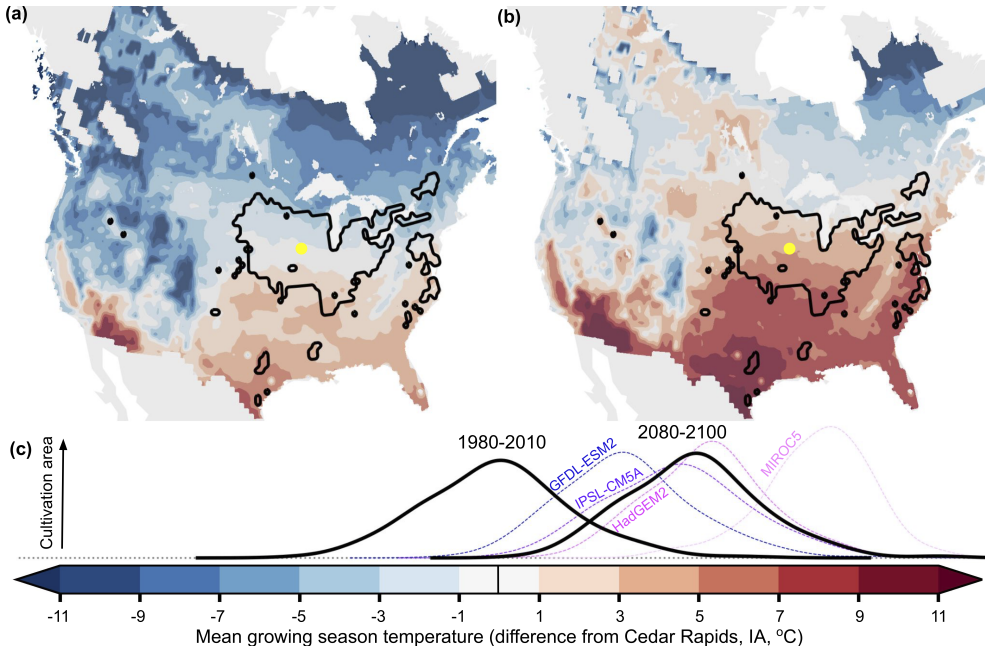


Figure 4.1: Poleward shift of optimum growing season temperatures under anthropogenic warming. Heat map shows the mean growing season temperature difference from the modal cultivated land for maize (Cedar Rapids, IA, for example - yellow dot) in **(a)** 1980-2010 **(b)** and the end of the century (2080-2100) under RCP8.5 for the InterSectoral Impact Model Intercomparison Project (ISIMIP-Fast track)-CMIP5 [69] climate model mean. Growing seasons are taken from the GGCMI phase 2 crop calendar for each gridcell and held constant under warming. Black contours in (a, b) show gridcells with at least 10 kha cultivated in North America. **(c)** Temperature difference distribution of all breadbasket maize cultivated areas in North America historical mean conditions (1980-2010) and the end of the century (2080-2100, RCP8.5). Optimum growing season temperatures in the center of the bread basket may move between 500km (GFDL-ESM2, low sensitivity) to 1200km (MIROC5, high sensitivity) northward by the end of the century under warming. Source: [49].

Plain is concentrated at similar temperatures. We therefore assume the Iowa growing period temperature is a de facto "optimum".

Expected warming in the corn belt by end of century is far larger than the current growing-season spread in temperatures. Using 4 representative models from ISIMIP [the InterSectoral Impact Model Intercomparison Project 69] and the RCP8.5 forcing scenario, the average temperature rise over these same growing periods is over 5°C (Figure 4.1, bottom). This increase is over 60% greater than the global mean temperature change (3.6°C) [99], as is expected for the continental mid-latitudes where the corn belt and other breadbaskets lie [e.g.

73]. Even in a strong emission-reduction scenario (CMIP6-SSP126), mean corn belt warming in these models is as high as ∼5°C [CMIP6 136, 194] by end of century (Supplemental Figure S51, S52). The high-end warming could translate into a poleward shift of maize cultivation by more than 1000 km, if the growing season "optimum" temperatures are to remain constant, i.e. a relocation of the current maize belt to central Canada (Figure 4.1, right).

4.3.2 Stasis in the No-Adaptation Case

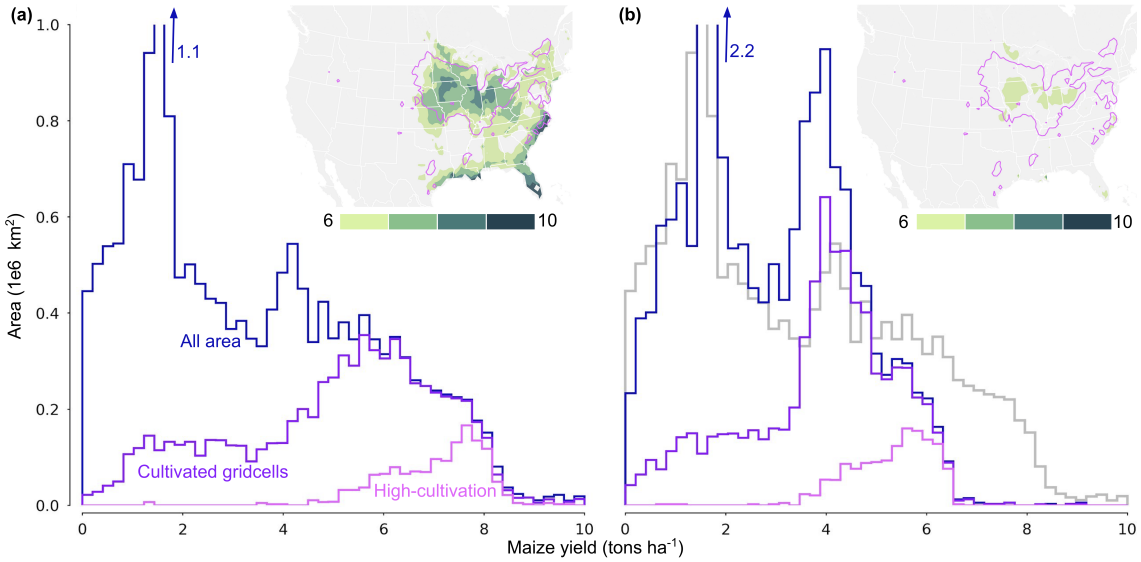


Figure 4.2: Breadbasket yields decline under warming while the latitude of peak yield remains fixed. Figure shows simulated yield across all land in North America in the GEPIC model for maize. Histogram lines indicate the median yield for all areas (blue), all areas with some cultivation (purple) and high-intensity cultivation areas (pink) for **(a)** 1980 - 2010 historical period and the **(b)** under 6°C of warming from the historical with constant [CO_2]. High-cultivation gridcells are those above 10 kha in North America. Inset plots show yields (in tons ha^{-1}) and the high-cultivation areas under the same conditions in each panel. Gray line in panel **(b)** shows the 1980-2010 distribution for all simulated areas. The highest simulated yields decline under warming and are not replaced at higher latitudes, therefore Iowa remains the best simulated maize yield in North America. See Supplemental Figures S53-66 for case with high [CO_2] and other models. Source: [49].

Without adaption, crop models generally show little poleward movement of optimal growing regions (gridcells in the top 10% of simulated yields, regionally) under warming. For North American maize, in the no-adaptation (A0) GGCMI Phase 2 6°C warming simula-

tions, mean poleward movement of the breadbasket center averages only 1.3 degrees latitude or less than 150 kilometers across 6 of the 7 models. (The pDSSAT model is an outlier both in temperature sensitivity and in changes in spatial patterns, and shows a shift of nearly 900 km.) All crop models show overall yield losses under warming, so that the maize breadbasket becomes less productive, but in all but pDSSAT, the current region of highest yield largely remains the optimum land. As an illustration we show simulation results from a single model, GEPIC (Figure 4.2). (See Supplemental Figures S53-S66 for other models, which show qualitatively similar results.) Under current climate conditions, simulated potential maize yields vary widely across North America given the diversity of climate and soils (Figure 4.2a, blue), with the highest potential yields in the real-world area of most intensive cultivation (Figure 4.2a, pink). Under 6°C warming, the yield distribution contracts and maize yields above 6 ton ha^{-1} almost entirely disappear. That is, yield declines across the "breadbasket" area are not paired with yield increases elsewhere, so the current Midwestern "breadbasket" remains the top-yielding region, although its productivity is lower. The slight net poleward movement of the model maize breadbasket (1.2 degrees or 133 km) is largely driven by more severe losses in the high-yielding Gulf Coast than in the Midwestern corn belt (Figure 4.2, insets), likely due to non-linear temperature impacts. (Losses in the Gulf Coast also contribute to the anomalous poleward breadbasket shift in pDSSAT; see Discussion and Supplemental Figure S56.)

The result that breadbaskets remain fixed in the absence of adaptation is robust across assumptions about $[CO_2]$ and water stress. Because $[CO_2]$ impacts are largely uniform across latitude, they can do little to modify the location of the "breadbasket" under warming. Using 810 ppm in the GGCMI Phase 2 +6°C warming simulations mitigates yield losses and can produce a slight westward shift of the breadbasket, presumably because elevated $[CO_2]$ protects against water stress (Supplemental Figure S53), but poleward shifts are even smaller than in the constant $[CO_2]$ case: only 1.1 degrees in the 6-model average. For pDSSAT,

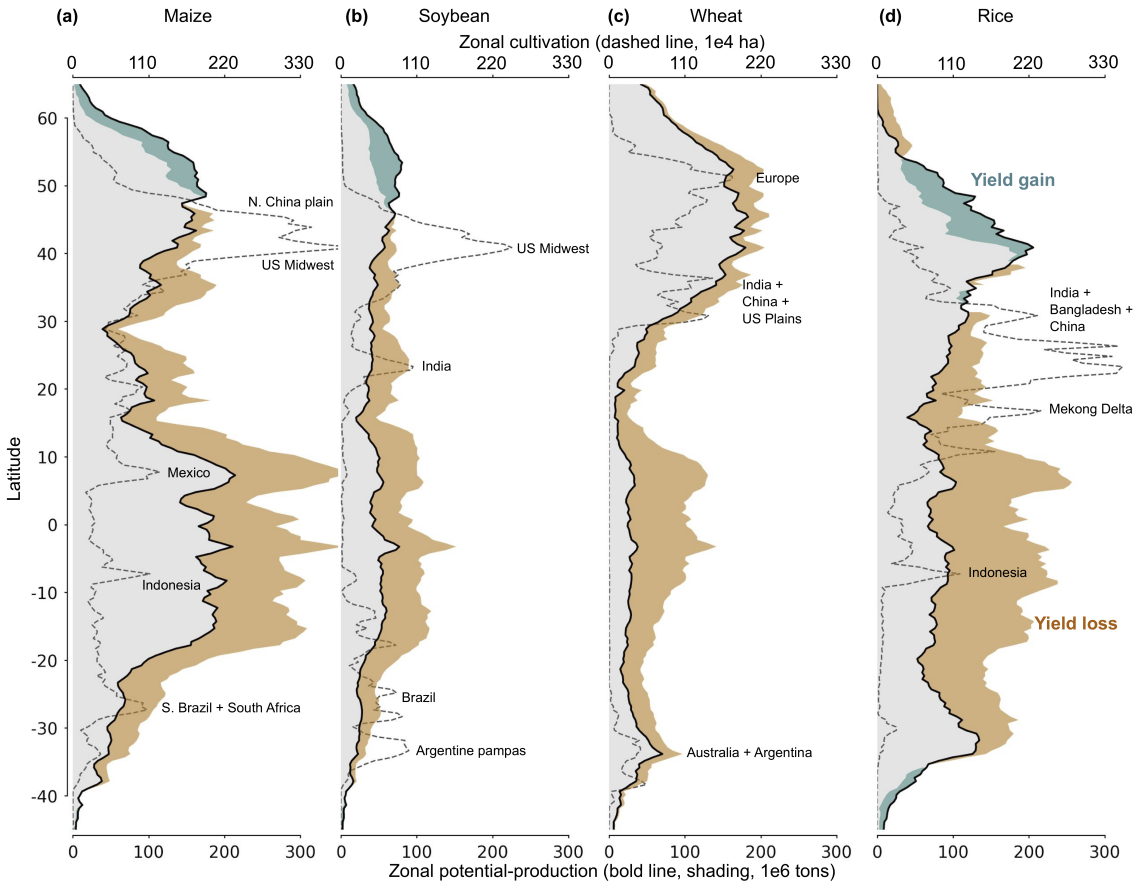


Figure 4.3: Warming-related changes in zonal potential grain production in the absence of adaptation, for maize **(a)**, soybean **(b)**, wheat **(c)**, and rice **(d)**, from GGCMI Phase 2 A0 crop model simulations. Values are total production in each 0.5 degree latitude band from all gridcells above global median yield, to approximate lands on which crops would reasonably be grown; values are calculated for each model and figure shows the multi-model mean. (See Supplemental Figure S68 for individual models.) Black line shows production under +6°C warming and brown and green areas the change from baseline climate conditions. Simulations are rainfed everywhere for all crops other than rice, which is irrigated everywhere. For reference, dashed line shows current actual total zonal cultivated area for each crop and labels mark regions of intense cultivation. Warming causes substantial losses at lower latitudes in all crops and some gains at high northern latitudes for all but wheat, but gains are not large enough to substantially shift the zones of optimal production at continental scale (e.g. N. American and Central American corn belts). Net production loss in wheat at high latitudes occurs because warming makes land unsuitable for high-producing winter wheat. (Wheat values here combine separate spring and winter simulations and take the better yield in each gridcell.) See Supplemental Figure S71 for simulations with elevated [CO_2], which show higher yields under warming but similar relative stasis of breadbaskets. Note that for crops other than rice, model peak production zones correspond to actual zones of intense cultivation, validating results. Real-world rice cultivation is constrained by availability of water and competition with higher-value maize and wheat. Source: [49].

the poleward shift is essentially identical for both [CO_2] levels (See Supplemental Figures S53–S66). Similarly, breadbasket location remains fixed in simulations without water stress (the GGCMI Phase 2 fully irrigated case) (Supplemental Figure S54). Unlimited irrigation produces a more westward breadbasket in all climate conditions, and mitigates yield losses under warming.

Results are also robust across crops and locations. Figure 4.3 shows an estimate of latitudinal-average total production for maize, soybean, wheat, and rice in both present-day and +6°C conditions. (To restrict production to plausibly cultivated land, we use only gridcells exceeding the global median for each model, then take the multi-model mean.) Warmer temperatures reduce production at most latitudes, with mild gains in northern regions for all crops other than wheat, which is further impacted by switching between winter and spring wheat. Northern gains are however not large enough to substantially modify the latitude of peak productivity. In the N. hemisphere, the poleward shift in peak productivity (total for a given latitude band) under +6°C warming is 2.5 degrees latitude (280 km) for maize, 0.4 (45 km) for wheat, 0.1 (10 km) for rice, and 3.6 (390 km) for soybean in the multi-model mean. (In the S. hemisphere, poleward motion of peak productivity regions is limited by lack of land). In simulations where warming is combined with high levels of [CO_2], the poleward shift in peak productivity under warming is even smaller: 2.3 degrees latitude for maize, 0.4 for wheat, -0.2 for rice and 3.2 for soybeans in the multi-model mean (Supplemental Figure S69, for the +6°C, 810 ppm [CO_2] case). Higher [CO_2] levels effectively eliminate warming-induced losses everywhere outside the tropics, but gains at high latitude remain too small to substantially shift zones of peak productivity.

4.3.3 Growing Season Adaptation Allows Geographic Shifts

Inclusion of adaptive measures in warmer climate conditions produces dramatic changes in projections of both yield amounts and their spatial distribution, in all models other than

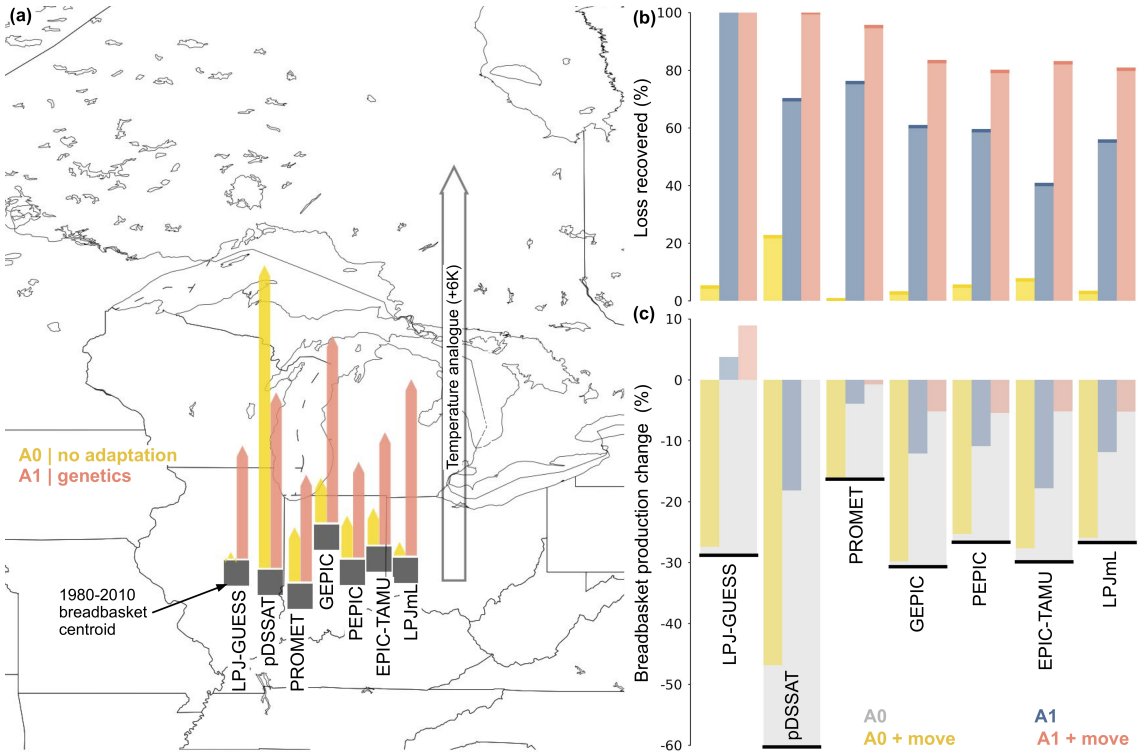


Figure 4.4: Effect of including cultivar adaptation on breadbasket location and productivity changes under warming, for N. American maize in the GGCMI Phase2 A0 (*no adaptation*) and A1 (*maturity delay*) simulations. Comparison shows baseline (1980-2010) climate and +6°C warming, [CO_2] of 360 ppm. Breadbasket are defined as described in text. **(a)** Mean latitude of the maize breadbasket for the 7 crop models and its poleward movement under warming for A0 (red) and A1 (yellow) simulations. Note that each model puts the historical breadbasket at a slightly different latitude, and longitudinal position has no meaning in this Figure.) All models other than pDSSAT, the high outlier in temperature response, show only slight poleward movement in A0 but substantial movement in A1. For comparison, white arrow marks the "climate analogue", the displacement in latitude required for the existing high-cultivation maize area to maintain its current growing season temperature. Even under A1, breadbaskets do not shift sufficiently to avoid warming. **(b)** Percent of temperature-driven production losses avoided by shifting cultivation locations alone (A0, yellow), by cultivar adaptation alone (A1, blue), and by both cultivars and movement (A1, red). Scale is truncated at 100%. Without cultivar adaptation, moving growing locations poleward has little benefit. Cultivar adaptation alone avoids 2/3 of warming-induced losses, and allowing relocation of cultivation further reduces losses. **(c)** Simulated maize production changes in the base +6°C warming case (grey = A0 + fixed cultivation) and under different adaptive measures (as in **b**). Changing both cultivars and locations reduces mean production losses from ~30% to <5%. See Supplemental Figures S70–71 for cases with +2°C warming and high [CO_2]. Source: [49].

pDSSAT. While crop model breadbaskets barely move under warming in the no-adaptation (A0) case (Figure 4.4a, yellow arrows), the same models produce significant shifts in the adaptation (A1) case (Figure 4.4a, red arrows). In the GGCMI Phase 2 simulations, all

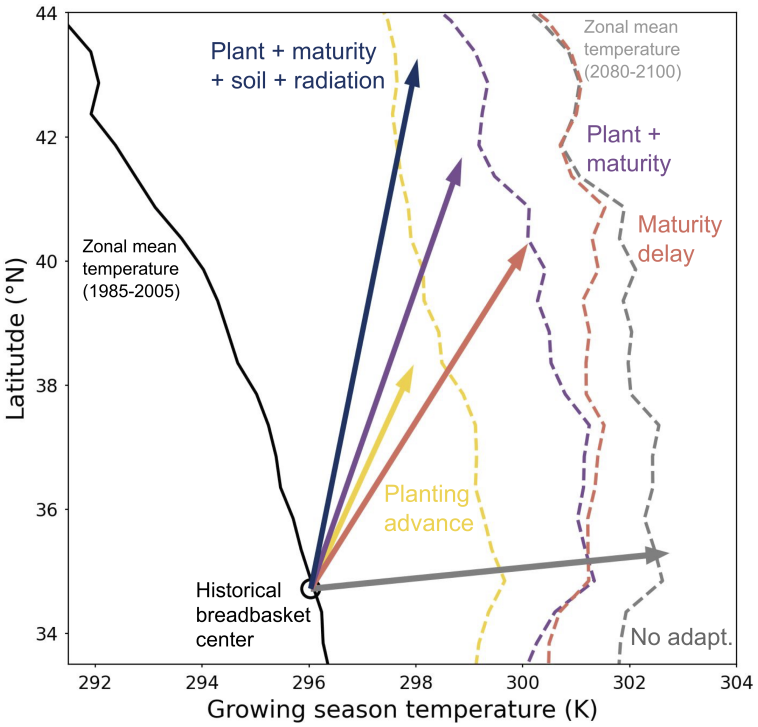


Figure 4.5: Changes in mean growing-season temperature and breadbasket location for N. American maize in LPJmL simulations with different adaptation strategies, as described in text. Arrows indicate breadbasket movement in temperature-latitude space by the end of the century in the HadGEM-2 RCP8.5 climate simulations. A vertical arrow would mean the optimal production location tracks with the temperature analogue. Dashed lines show zonal mean temperature rise at each 0.5 degree latitude band, for all N. American gridcells East of 100W. Note that climate projections flatten the meridional temperature gradient. Simulations show here use fixed [CO_2] for consistency; results are nearly identical in the dynamic [CO_2] case (Figure S72). With no adaptation (grey), breadbasket location remains nearly constant. All growing-season adaptations allow more substantial movement. Allowing earlier planting (olive) produces a moderate shift in breadbasket location and a much lower temperature rise, since early planting and accelerated maturity mean harvest occurs in mid-July instead of early September. Allowing delayed maturity (green, teal, blue) produces much larger breadbasket shifts but also warmer growing-season temperatures. Removing soil and radiation constraints (dark blue) contributes to poleward motion, even in this case total breadbasket shift is about half that required to produce a temperature analogue to present-day conditions. Source: [49].

models have a harmonized growing season of ∼135 days under current conditions for the N. American maize breadbasket (ranging from ∼130 days in Southern Illinois to ∼140 days in North Dakota). In the A0 simulations, the growing season contracts in warmer conditions in all models, losing an average of 33 days at T+6 C, or 25% of its total length. Shortened growing seasons under warming result because plant maturation is tightly related to total

accumulated heat, and contribute to steep yield losses averaging 31% over the breadbasket. Because yields decline everywhere, moving maize cultivation slightly north to the new optimal location can recover only a few percent of warming-related losses (Figure 4.4b-c, yellow bars). In the GGCMI Phase 2 A1 simulations, model parameters are adjusted to retain growing season length, representing adaptation by using new cultivars with delayed maturity. Cultivar adaptation alone recovers on average 68% of the temperature-driven loss seen in the A0 case, even if cultivation area is held fixed (Figure 4.4b-c, blue bars; see also [216]). Because areas further north actually rise in yield, the optimal location for maize cultivation in North America shifts north by on average 3.1 degrees latitude (340 km) excluding pDSSAT, or 3.3 degrees (370 km) across all models. Allowing maize cultivation to move north to its new optimum location further reduces losses related to climate change, by an average of 91% across models. These results are qualitatively similar across degrees of warming (Supplemental Figure S70 shows the +2°C case), and whether or not [CO_2] is assumed to rise with warming. When warming of +6°C is combined with [CO_2] rising to 810 ppm (Supplemental Figure S71), the breadbasket shift in cultivar-adaptation A1 scenarios is 2.4 degrees (266 km), vs. 1.1 degrees (120 km) in the no-adaptation A0 runs (excluding pDSSAT).

Although breadbaskets do move substantially in the GGCMI adaptation (A1) simulations, they still shift less than would be expected in a "climate suitability" framework. In the +6°C applied warming scenario, the mean N. American maize breadbasket would have to shift by over 800 km to maintain a constant growing-season temperature (Figure 4.4a, white arrow). The realized shift of 340 km in the A1 case (for all models excluding pDSSAT) mitigates only 3.5°C of warming. (In the temperature-sensitive pDSSAT, its larger geographic shift mitigates all but 0.6°C of warming.) One potential explanation is that the A1 simulations only partially capture farmer adaptation. In A1, delayed-maturity cultivars allow retaining present growing-season length in each location, but in reality farmers would *extend*

growing seasons to take advantage of earlier frost-free conditions. The A1 specifications artificially leave planting and harvest dates unchanged even in warmer conditions. We therefore turn to new simulations with the LPJmL model to explore the full set of factors that may affect changes in spatial yield distributions under warming.

The LPJmL model allows incorporating rule-based growing season adaptions in addition to a crop maturity delay. We consider an extended set of simulations including four adaptation cases. In *planting advance*, planting dates are allowed to advance with warming but cultivar choice is unchanged. In *maturity delay*, cultivar adaptation maintains growing period length but planting date is unchanged, analogous to A1. In *plant & maturity*, both adaptations, and, for comparison, in *soils & radiation* full adaptation is combined with uniform soils and radiation set to values in Iowa, the presumed optimum. Simulations over present-day and warmer future conditions are run using output from 4 climate models as described in Methods; we show HadGEM2 results here, in which the mean future warming of the historical N. American maize breadbasket and growing season is 6.7°C. As expected, the LPJmL runs are consistent with the conclusion that in the absence of growing season adaptations, optimal growing locations do not shift (Figure 4.5). In the LPJmL *no-adaptation* case under HadGEM2, the N. American maize breadbasket moves only slightly poleward, by 55 km, and experiences nearly the full extent of unmitigated warming, 5.25°C out of 5.28°C in the original location (Figure 4.5, grey), and severe yield losses at -20%. Note that this temperature differs from the 6.7°C given above because the growing season contracts dramatically, by more than 5 weeks or ∼30%, so that harvest occurs in late July instead of early September. See Supplemental Table S3 for growing season details and yield losses for all cases.

Growing season adaptations appear critical in allowing the optimal maize production location to shift under warming, with geographic factors playing a lesser role. Allowing *maturity delay* adaptation produces a strong poleward shift by nearly 610 km, larger than

in the analogous GGCMI Phase 2 A1 simulations (Figure 4.5, green), but only weakly mitigates warming, since in these more realistic climate simulations the meridional temperature gradient flattens. Because high latitudes warm more strongly than mid-latitudes, displacing cultivation poleward is less effective at lowering growing season temperatures than in GGCMI Phase 2 where temperature offsets are uniform. Yield losses are then nearly as great as in the *no-adaptation* case, at -15%. In the *planting advance* case, the poleward shift is more moderate, at 400 km (Figure 4.5, olive). Growing season temperature rise is however kept to only 1.9°C by a combination of severe advancement of the growing season (by three weeks, so that planting occurs in early April instead of early May) and contraction of the growing season (by over a month, so that harvest now occurs in early July). Yield losses remain severe at -15%.

Substantial avoidance of yield losses occurs only when both growing season adaptation strategies are adopted simultaneously, so that warming can be reduced while maintaining growing season length. In the *plant & maturity* case (Figure 4.5, teal), peak yields move poleward by 780 km, planting dates remain advanced, and mean growing season becomes actually longer than in the baseline case, so that harvest occurs in late August. The longer time to maturity means the growing-season mean temperature is warmer than in *planting advance*, at 2.8°C, but yields are substantially higher, at net gains of 3% from the baseline case. Finally, removing constraints imposed by soils and radiation boosts the poleward breadbasket shift and yield gain only slightly, suggesting these are not the major constraints (Figure 4.5, blue). Using Iowa *soils & radiation* conditions everywhere moves the optimum production region an additional 130 km polewards, reduces net warming to 2.3°C, and increases yields to 6% above baseline.

Note that even the maximum breadbasket shift in Figure 4.5 is still only about half that required to produce a "temperature analogue" to present-day conditions, as in the climate suitability framework. It may in fact be geographically impossible for cultivation

to shift sufficiently to avoid all warming, since in these realistic climate projections the temperature analogue moves twice as far as in the GGCMI Phase 2 simulations of Figure 4.4. Nevertheless, growing season adaptations in LPJmL are sufficient to retain yields. These results are qualitatively similar for all 4 climate models tested, even though amounts and patterns of warming differ (Figure S72). While Figure 4.5 shows the fixed-[CO_2] simulations, shifts are nearly identical in the dynamic-[CO_2] case, although yields improve (Figure S72, see also Figures S26–30 for spatial distributions of LPJmL yields in each climate simulation).

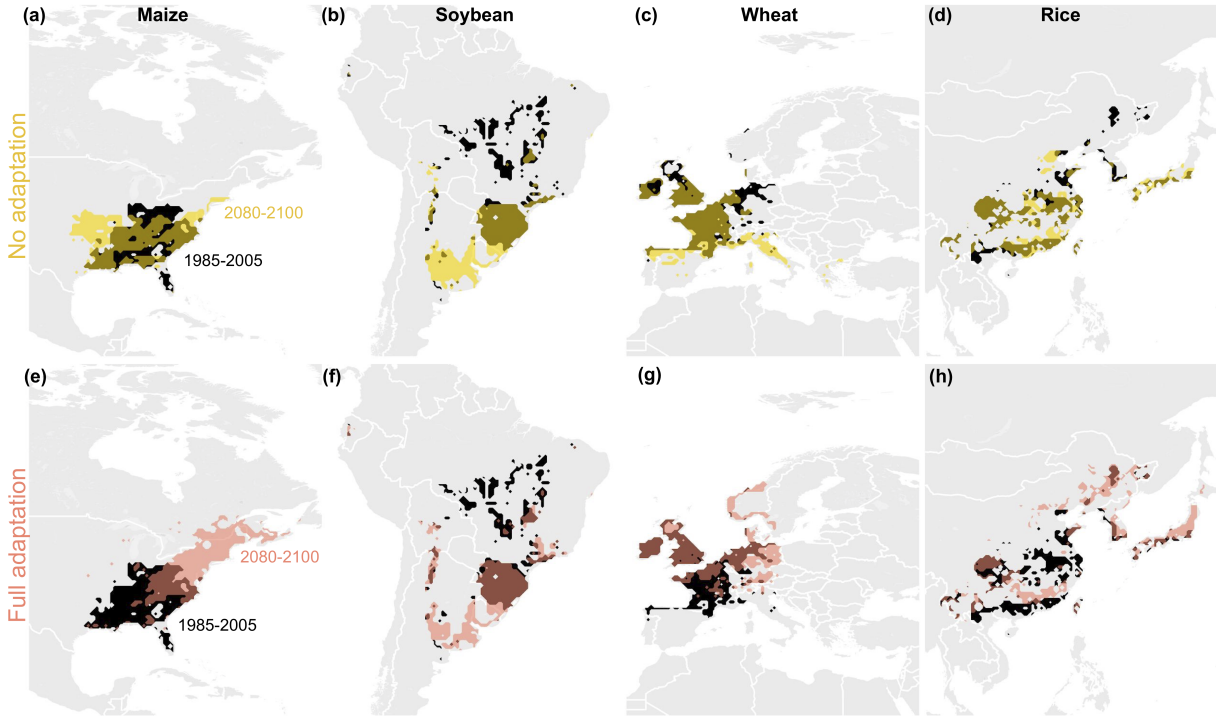


Figure 4.6: Regional differences in breadbasket shift. **(a-d)** historical top 10% regional yields (black) and end of the century top 10% of yields (gold) with fixed planting days and cultivar genetics for four crops and regions. **(e-h)** Same as top row now with fully-adapted growing seasons (planting advance and maturity delay). Simulations from LPJmL driven by the HadGEM-2 climate model under RCP8.5. South American soybeans move poleward to a further extent with fixed growing seasons than other crops. Wheat in Europe and Rice in East Asia both move poleward to a greater extent with dynamic seasons, but neither to the extent of North American Maize. See Section 2.2 for discussion about historical yield patterns. Source: [49].

These results are broadly consistent across other major global breadbaskets. Figure 4.6 highlights the shift of the top breadbasket for each of four crops included in this analysis un-

der both *no adaptation* (top) and the fully adapted *plant & maturity* (bottom) simulations. In all crops other than soybean, breadbaskets remain almost stationary under warming in the absence of growing season adaptation, or even move slightly equatorwards. The exception is S. American soybean, whose optimal production location moves strongly poleward, but this case is also the exception in that the breadbasket lies in the subtropics (21 S) rather than mid-latitudes, where precipitation rather than temperature typically drives the spatial distribution of yields. With full growing season adaptation, all breadbaskets move moderately to strongly polewards. Both maize and soybean breadbaskets shift by nearly 800 km. Rice cultivation in Asia is constrained by land availability, since simulations were run only over areas where irrigation is feasible (excluding deserts), and moves only ∼500 km. European wheat exhibits the weakest shift, ∼350 km, again in part from land considerations: wheat cultivation is already centered so far north (50 N) that limited lands exists for further poleward displacement (Figure 4.6c,g). See Supplemental Table S4 for summary statistics for all crops.

4.4 Discussion

In experiments with multiple crop models, inclusion or omission of adaptive measures substantially impacts the location of optimal grain production. Breadbaskets do not move poleward under warming if changes in planting date and cultivar genetics are not permitted, but do shift substantially given these adaptations, with the largest response resulting from cultivar changes to delay maturity. With full growing season adaptation, the projected shift in the North American maize belt by end of century under the RCP8.5 scenario is almost 800 km (approximately 10x the shift observed since 1973 in [184]). Yields are recovered only in full adaptation simulations that allow both adopting new genetics and adjusting sowing dates: both adaptations appear required to fully capitalize on the temperature increase at higher latitudes. These results are common to nearly all crop models and are robust to the

level of warming, whether simulations apply uniform warming or more realistic patterns, and whether [CO_2] fertilization is or is not considered. As expected shifts appear slightly greater in cases with uniform warming (as in the GGCMI experiment) than in those with more 'realistic' climate scenarios with amplified high-latitude warming. In no cases, however, do breadbaskets shift sufficiently to follow the temperature analog, even when constraints of soils and radiation are removed. Water availability could play some role in limiting poleward shifts, but note that *changes* in precipitation do not, since the four climate models tested do not show consistent latitudinal differences in precipitation changes. (See Supplemental Section S3.1 for details.)

Although the crop models used in this work produce qualitatively similar results, it is not yet possible to quantitatively predict future breadbasket locations. While the inclusion of adaptation in crop models is an advance on assuming static growing seasons, as is often done in global-scale crop model experiments [e.g. 165], all versions shown here are still simplified. In LPJmL, the rules used to determine planting dates and cultivars use only one of either temperature or precipitation, depending on mean seasonal characteristics [204, 120], and neglect considerations of intra-annual variability as well as soil moisture dynamics that can constrain planting dates. Furthermore, while planting dates are adjusted in each individual year, cultivar choice is based on 20-year mean climate and is not adjusted in years where planting is earlier or later. The model does not take account of management issues like the need for double-cropping, or potential limitations in the availability of ideal cultivars for future climate conditions [e.g. 216]. In general, many crop models do not have a sophisticated enough soil representation to capture all limitations imposed by poor soil quality, as in the thin-soiled Laurentian Shield of the Upper Midwest and Canada. Finally, the 50-km horizontal resolution of this study may not be suitable in cases where agriculture moves to higher *altitudes* under warming, a potentially important adaptation pathway [e.g 102].

Crop models also show intrinsic differences in the growing season responses to warming

that drive the need for adaptations. The GGCMI Phase 2 crop simulations differ in the contraction of growing seasons in the no-adaptation case, and, especially at high latitudes, in the ability of cultivar changes to recover original growing seasons. In general, models show the strongest agreement in currently-cultivated areas, and the largest spread in the high-latitude areas where breadbaskets may move (Supplemental Figure S79). Treatment of model calibration may also be a source of differences in model projections of breadbasket shifts. While most GGCMI Phase 2 models are uniformly uncalibrated, the outlier in this analysis, pDSSAT, was calibrated in those regions where yield data exists. This difference may contribute to the stronger temperature sensitivity of pDSSAT maize [48], and in turn in the greater poleward displacement of its breadbasket.

While this study focuses on changes in potential yields, it is also worth noting that the spatial distribution of agriculture today depends on many other factors. In many places, food is grown where people live, meaning that much current cultivation occurs in sub-optimum conditions [e.g. 131]. Areas of intense cultivation also often relate to national borders, whether because of cultural factors or through strategic choices to maintain domestic production. (See e.g. local cultivation peaks in India and Indonesia in Figure 4.3.) Simulated shifts in breadbaskets across national borders may not materialize for the same reasons that have established patterns of cultivated areas today. (See Supplemental Table S3 for transnational shifts in LPJML.) Even within individual countries, production of many crops are physically constrained by geography: for example, rice in Indonesia, wheat in Southern Australia, or any crop in Northern India or South Africa. (See Supplemental Figure S78 for shifts in African breadbaskets.) Finally, cultivation of one crop can affect others. For example, soybean is often planted in areas that may be warmer than biologically optimal, because it is rotated with maize for restoring soil nutrients. All these considerations mean that adaptation by shifting cultivation may be limited – that is, in future warmer conditions, cultivation may not shift as much as do potential yields.

The results of this study nevertheless suggest important lessons for understanding the effect of climate change on food production. Agricultural climate impacts assessments to date have typically focused on currently-cultivated land and generally omit adaptive measures. This is necessarily true for studies based on statistical models [e.g. 180, 77], and those using process-based models have largely followed similar practices [e.g. 165, 123, 88]. More recently, some studies have included cultivar adaptations [e.g. 6], but do not include changes in planting dates. Current assessments are therefore likely overly pessimistic, both overestimating damages where crops are currently grown and not allowing shifting cultivation to higher-yielding areas. These issues in turn compromise economic studies that build on these assessments to estimate changes in regional agricultural competitiveness [e.g. 189, 91]. This work highlights the complexity of capturing agricultural responses to climate change, and the limitations of "shortcuts" in analysis: neither assumptions of stasis nor simple climate suitability frameworks appear to apply. Furthermore, understanding the displacement of breadbaskets is intrinsically important, because the climate threat to the food system lies not only in reductions in total food production, but in disruptions to where that production happens. If adaptive measures substantially change both yields and spatial yield patterns under climate change, accurately capturing agricultural responses requires global analyses using process-based models that explicitly represent these management choices. Such efforts are more difficult, and may require large interdisciplinary teams, but appear to be the only pathway to understanding climate risks to the global food system.

4.5 Data Availability

All data used in this analysis are publicly available. Simulation outputs from GGCMI phase 2 are available on at *https://zenodo.org/*. CMIP5 climate model outputs are publicly available at ISIMIP *https://data.isimip.org/*. AgMERRA data is available at reference: [171]. SPAM data is available at reference: [84]. LPJmL RCP8.5 simulation outputs and code for analysis

and figures are available upon request to the authors.

CHAPTER 5

NUTRIENT CONTROLS ON CLIMATE RESPONSE AND UNCERTAINTY FOR MAJOR GRAINS

5.1 Introduction

The global community faces the large challenge of meeting rising food demand in the next century under the pressures of climate change. Nitrogen fertilizers have played the predominant [186, 37, 190, 182, 125] role in the tripling of global vegetable proteins [104] in the last half-century on relatively constant agricultural land. Without synthetic nitrogen fertilizers, the global food system could not secure the diets of as many as 3.5 billion people alive today [37, 185]. While models project yield damages due to climate change by the end of this century [e.g. 180, 165, 178], large uncertainty remains [22, 5] including the interaction between applied nitrogen fertilizers and climate change induced abiotic stresses on crop production at the global scale. In this study we use an ensemble of six globally-gridded-process-based crop models for wheat maize, and rice with simulations across a range of weather and management input variables to study the interaction between weather, [CO_2], and nitrogen fertilizers in order to test the potential for future fertilizer intensification under climate change. Here we show that available nitrogen levels in soil strongly determine climate change impacts on global grain yields and that nitrogen availability causes yield losses for the major grains to increase from 6% per K to 10% per K up to 6 degrees of warming in the global aggregate on currently cultivated land.

Unabated climate change will likely have a negative impact [e.g. 180, 165, 178] on most major crop yields during the 21st century, although studies differ on the magnitude and sometimes the direction of the impact [22]. Absolute reductions in global production will likely not be realized however, because production needs to double by 2050 in order to keep pace with population growth and changing tastes [57, 196]. Presently over 40% of

the land surface –including deserts and Antarctica– is used for food production (crops and livestock grazing) [46]. While some cropland expansion is to be expected (especially in the northern latitudes), much of the required production increase must come through sustainable intensification on currently cultivated lands if we are to preserve remaining forests and other natural habitats.

Applied nitrogen fertilizers have likely played the most critical role in the 3x increase in vegetable protein production globally in the last 60 years on relatively constant agricultural lands [186, 37, 190, 182, 125]. Nitrogen fertilizer intensification will continue to play a critical role –especially in developing regions such as sub-Saharan Africa, Southeast Asia, and Latin America– in meeting rising food demand alongside the pressures of climate change. The numerical models necessary to project the future impacts of climate change on agriculture often do not represent the details of the nitrogen process in plant growth, and when they do, only simulate a static nitrogen application level. The recent addition of the nitrogen cycle to dynamic vegetation models has shown to provide an increase in model performance [203] and has shown to predict steeper reductions in yield under climate change [165, 11]. However, no study has investigated in detail dynamic nitrogen inputs alongside changing climatological inputs at the global scale for the major grains. Here we present the first global-high-resolution modeling study to systematically test different climate inputs on various fertilizer application levels for the three major grains. The Globally-Gridded Crop Model Intercomparison (GGCMI) project allows testing the robustness of this relationship for the first time at the global scale across many different process based crop models.

Agriculture's central objective is the production of human-digestible nitrogen [106]. Plants utilize solar energy to combine water and nutrients from the ground and carbon from the air to make chains of organic molecules. The species of plant that dominate the global food production system are exceptionally skilled at utilizing nitrogen to form amino acids, the building blocks of life for all eukaryotes. In the natural nitrogen cycle, bacteria in the soil fix

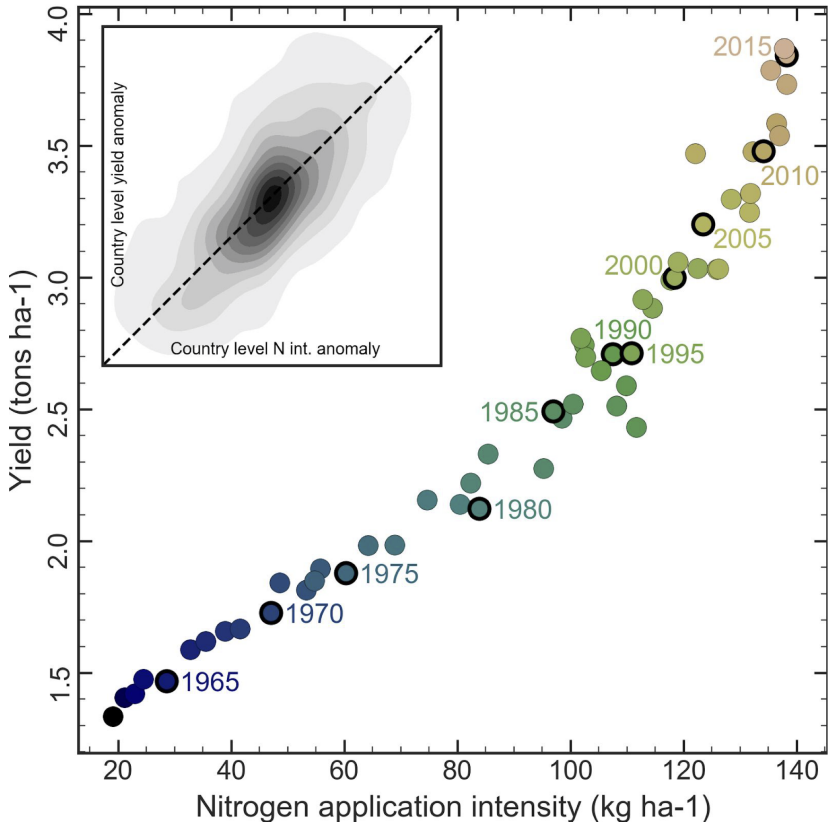


Figure 5.1: Historical nitrogen intensity and yields. Global mean nitrogen application intensity vs global mean yield for the 10 most cultivated crops globally (maize, rice, wheat, sorghum, barley, millet, seed cotton, and groundnuts; with soybeans and beans held out)[46, 83]. Inset: kernel density plot of country level nitrogen application intensity anomaly vs. yield anomaly for the same crops across most countries individually. Global mean yield of the top producing crops has increased 2.2x and nitrogen application intensity has increased 7x in this period. Global production has increased threefold in total and fertilizer applications have increased 10x since 1960. Non-representative small island and desert nations removed including: Mauritius, Qatar, Kuwait, Cyprus etc. Note: inset plot clipped y-axis to ±2 ton ha-1 and x-axis to ±100 kg ha-1 for clarity. (See also supplemental Figure S80.)

atmospheric nitrogen (N2) and water to form mineral nitrogen (e.g. NH_3) which can then be directly utilized by the plants. Since the early 20th century, when Fritz Haber developed a process for artificially manufacturing NH3 from its component, synthetic nitrogen fertilizers have dominated the changing agricultural landscape. Fertilizer applications have increased 10-fold since the 1960's [83] and global yields have increased 3-fold (Figure 5.1). The global food system could not secure the diets of as many as 3.5 billion people alive today [37] without synthetic nitrogen fertilizers.

Plants take up nitrogen from the soil at a rate determined by soil moisture and photosynthetic efficiency and metabolize it into proteins at a rate determined by environmental conditions and biotic stresses during the grain-filling stage of growth. In this study, we utilize six different process based crop simulation models that parametrize the nitrogen uptake and metabolism processes via several different approaches. The models are run across a variable space of potential future climate scenarios that vary temperature, precipitation, and atmospheric [CO_2] at different nitrogen application levels for the three major grains globally; wheat, maize, and rice.

5.2 Methods

Overview of global gridded crop models (GGCMs). Model simulations used in this analysis were compiled as part of GGCMI Phase II under the AgMIP umbrella cite [163]. See [47, 48] and citations within for more details. The models included are all mechanistic process-based crop models that are widely used in impacts assessments and represent over 3 decades of model development. Each model is run at 0.5 degree spatial resolution and covers all currently cultivated areas and much of the uncultivated land area. The six GGCMs are:

- the parallel Decision Support System for Agro-technology Transfer (pDSSAT) model [36, 94]
- the Lund-Potsdam-Jena managed Land (LPJmL) model [203]
- the Environmental Policy Integrated Climate: Texas A&M (EPIC-TAMU) model [86]
- the Geographic Information System-based Environmental Policy Integrated Climate (GEPIC) model [109, 42]
- the Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS) [107, 135]
- the Python-based Environmental Policy Integrated Climate (PEPIC) model [110, 111]

Site-Base models and Ecosystem models. Models fall into one of two categories: site-based, and ecosystem models. Site-based models are extended in this case for global assessment with a geographic information system (pDSSAT, EPIC-TAMU, GEPIC, PEPIC). Ecosystem models were originally developed to represent the terrestrial biome carbon cycle and have since been expanded to simulated managed land (LPJmL, LPJ-GUESS). Each type has its benefits and drawbacks, with site based models typically containing more detail on cropping systems and ecosystem models more details on the water cycle for example. Given differences in output yields, similar genealogy models (e.g. EPIC family) are considered independent models in this analysis.

Modeling protocol. All GGCMs simulated maize, rice, winter wheat, and spring wheat expect LPJ-GUESS, which does not simulate rice. Nitrogen is applied in two doses: 50% at planting and 50% on a crop- and grid-cell-specific day. All other stresses except temperature- and water-related factors and those related to nitrogen and [CO_2] (e.g. alkalinity, salinity, pests) are disabled. Models make different assumptions about residue management between growing seasons. Planting and harvest dates are harmonized across models based on data from [157, 145]. The experimental protocol consists of 9 levels for precipitation, 7 for temperature, 4 for [CO_2] and 3 for applied nitrogen, for a total of 672 simulations for rain-fed agriculture and an additional 84 for irrigated crops. All impacts calculations are masked and weighted by the MICRA [145] growing area dataset for each crop. Impacts are calculated at the 30 year temporal aggregation.

Model representation of biophysical processes. Biophysical processes represented by the models include photosynthetic light utilization, soil water and nutrient dynamics, abiotic stresses and [CO_2] effects. Radiation use efficiency is a simple RUE parameterization or a detailed leaf-level photosynthesis respiration. Respiration of [CO_2] and transpiration rates follow either the descriptive or explanatory approach. In the RUE type models, increased [CO_2] proportionately affects the rate of photosynthesis. Transpiration response to varied

[CO_2] is different across models. More details on [CO_2] response can be found in reference29. All GGCMs simulate heat and water stress. Water demand and evapotranspiration in the EPIC models is based on the Penman-Monteith41 approach and the Priestly-Taylor [149] approach elsewise. Crop phenology in GGCMs is based on growing-degree day heat accumulation with crop-specific values for temperature (see Supplementary Table 3 in [165]).

Model representation of nitrogen use. Synthetic and natural fertilizers sources are not treated differently in the simulations– all fertilizers are treated as their synthetic equivalent. We do not consider any additional nitrogen inputs such as atmospheric-soil deposition. Different models make different assumptions on soil organic matter that may release additional nitrogen through mineralization. Biological N fixation only applies to soybean (not analyzed here), which always gets all the N it wants at no cost. The net mineralization (gross mineralization - immobilization) plus the prescribed N fertilizer input. The net mineralization only applies to the growing season of the crop so the mineralization of N during fallow periods is not recorded explicitly but still affects the N availability to the plants at the start of the season.

Nitrogen use by crops in process-based models is a parameterized balance between demand from photosynthetic processes and nutrient uptake based on soil and root properties. The nitrogen cycle and resultant nitrogen stress is calculated in the models. The specific nitrogen process is handled differently in each model, for example: LPJmL - daily photosynthesis and maximum carboxylation capacity (Vmax) are computed based on absorbed photosynthetically active radiation. The water stressed Vmax determines the demand for nitrogen by the plant. Uptake is determined by soil mineral N concentrations, fine root mass, soil temperature and porosity, and plant demand for N. This is computed for all soil layers individually and summed up to compute overall N uptake. Nitrogen allocation follows carbon allocation with functional and allometric rules and set limits on C-N ratios. See [203] for full details. LPJ-GUESS - Nitrogen uptake is computed on a daily time step and

is based on optimum leaf demand and ability for relocation from the leaves to maximize biomass growth. The pool of available nitrogen to plants is a result of gross mineralization and biological fixation by microbes, which is in turn a function of evapotranspiration based on N and C soil dynamics and are jointly simulated based on an SOM scheme adapted from the CENTURY model. See [187] for full details.

Climate inputs. Historical daily climate inputs are taken from the 0.5 spatial degree NASA AgMERRA daily gridded reanalysis product specifically designed for agricultural modeling, with satellite-corrected precipitation [171]. Emulated climatological-mean values are calculated from CMIP-5 [192] ensemble of models within the growing season for each crop in each location at the 30-year time scale anomaly between 1980-2010 and 1970-2099. $[CO_2]$ values at end of century are taken from [160].

Crop model calibration. Uniform, formal calibration of models is not performed as part of the model intercomparison. Site-based GGCMs are calibrated through a large set of parameters based on existing site studies. Ecosystem models are calibrated at the national level based on the FAO data by tuning a preselected small set of parameters. We always compare relative changes in yield across models.

Model simulated yield validation. The models are scored by comparison to the historical (1980-2010) data from the FAO at the country level. The models are also evaluated at the 0.5 degree grid cell level with comparison to the data from [157]. We calculate a pearson moment correlation coefficient between the detrended time series of yields from each data set and the simulated output yields.

Model emulation. We fit a third-order polynomial with interaction terms to the 30-year mean climatological yield simulation outputs. We use Bayesian Ridge probabilistic estimator is used to fit the parameters because it allow holding the functional form constant across space and across different models and provides stability in parameter estimates. The emulator captures the response of the simulation at the climatological mean level extremely

well, with mean absolute error weighted by cultivation no higher than 0.05% of baseline yield in the globally weighted average. We use the model emulator to complete the input dimension interaction space and calculate the impacts shown in this study unless explicitly stated as a 'simulation' output.

Nitrogen fraction of yield damages. The relative importance of the interaction between nitrogen and temperature under climate change is tested in the following way. We compare the difference in yield at present climate and nitrogen application ($Y_{currentT,currentN}$) from yield at present climate with no applied nitrogen ($Y_{currentT,noN}$) to the difference in yield at the end of century business as usual climate and current nitrogen application ($Y_{hot,currentN}$) from yield at difference in yield at the end of century business as usual climate and no applied nitrogen ($Y_{hot,noN}$). We then normalize this difference by dividing by the overall change in yield between the present and future at current real-world nitrogen application rates.

$$N_{factor}[\%] = \frac{(Y_{currentT,currentN} - Y_{currentT,noN}) - (Y_{hot,currentN} - Y_{hot,noN})}{Y_{currentT,currentN} - Y_{hot,currentN}} \quad (5.1)$$

CO_2 boost. The missed future [CO_2] direct fertilization boots at present nitrogen application levels is calculated using the model emulators. Yields at end of century under RCP 8.5 are emulated at both low (360 ppm) and high (800 ppm) [CO_2] at both present nitrogen application levels and 'full' nitrogen application levels. Impacts are aggregated up to regional level based on cultivation area from the MICRA2000 [145] dataset at temperature and precipitation projected at end of the century.

$$C_{boost}[\%] = \frac{(Y_{highC,highN} - Y_{lowC,highN}) - (Y_{highC,currentN} - Y_{lowC,currentN})}{Y_{highC,highN} - Y_{lowC,highN}} \quad (5.2)$$

5.3 Results

Crop simulation models parameterize biotic and abiotic stresses using a variety of methods and typically order the stresses in a hierarchy of varying degrees of formality (see methods). Under this modeling framework, the plant growth is limited by the most to the dominant stressor to varying degrees. In the EPIC models for example, the most severe stress dominates almost entirely. Simulated crops do not respond to stresses in a multiplicative sense in most cases. Instead, crops are less sensitive to temperature under nitrogen stress (not enough mineral nitrogen available) than are crops with full nitrogen availability and vice versa [85]. The simulated yield curves for crops at high and low nitrogen application levels for historical (1980-2010) simulations illustrate this stress hierarchy for the major grains (Figure 5.2). Nutrient limited crops show a muted response to the range of growing season temperatures when compared to the high fertilizer crops. Crop yields with and without nitrogen stress differ greatly at optimum growing season temperatures and become more similar in warmer growing seasons. For example, maize yields are indistinguishable above seasons with more than 2000 GDD, where the yield is very low in both cases, and therefore no additional benefit is gained from applied fertilizer.

This effect is further quantified for maize in North America with an econometric-style panel model where log simulated yields are regressed against GDDs and HDDs while accounting for local fixed effects. Again, we find that GDDs are more beneficial for crops without nutrients limitations and, importantly, HDDs are more damaging to yield (Figure 5.3). The impact of a day on yield at 26 C is slightly more beneficial to crops with high nitrogen inputs than to those with low. Conversely, a day at 40 C is dramatically more detrimental to yields (Figure 5.3).

In future climate change simulations, median fractional yield losses (from current production levels) due to climate change are amplified at high nitrogen by nearly a factor of two across all Koppen-Geiger climate regions for maize (supplemental Figure S81) and rice,

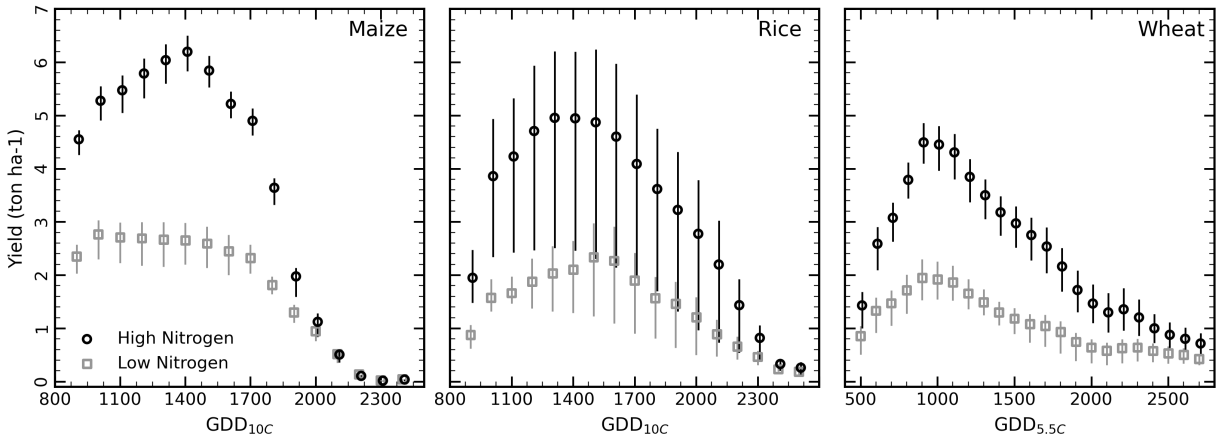


Figure 5.2: Historical simulation temperature-yield curves for maize, rice, and wheat. Simulation yield for rainfed maize (six models), irrigated rice (five models) and rainfed+irrigated wheat (six models) for the historical (1980-2010) period. Points show the mean yield for all grid cells with over 500 ha cultivated globally, binned by growing degree days in the last three months before harvest (bin size: 100 GDD). Error bars show confidence intervals calculated by holding individual models out of the mean calculation. Black shows 'full' nitrogen (200 kg ha-1 applied N) yields and gray shows nitrogen limited (10 kg ha-1) yields.

and across all but the cold continental region for wheat. High nitrogen availability results in higher fractional damages in almost every case and typically lower fractional increases for wheat in those regions where the elevated [CO_2] effect is driving increases in wheat production in the future. In some regions and for some crops, the nitrogen fertilizer level can change the sign of the yield changes for mid-century levels of projected warming. The downstream effect of this impact amplification is an potential mis-estimation of climate impacts of a given temperature change for regions which do not currently have high nitrogen inputs (Supplemental Figure S80). In the tropics, yield loss estimations may be 25% more optimistic than realized losses in a future hotter, more nitrogen-intense world (Figure 5.4).

The impacts of another critical climate variable, atmospheric [CO_2], are also determined in part by nitrogen fertilizer levels. The potential beneficial effect of elevated [CO_2] on yields is limited by nitrogen availability. We simulate the effect of increased [CO_2] on grain yields at the end of this century based on business as usual emissions paths at present nitrogen application levels. We calculate the difference in production over currently cultivated hectares

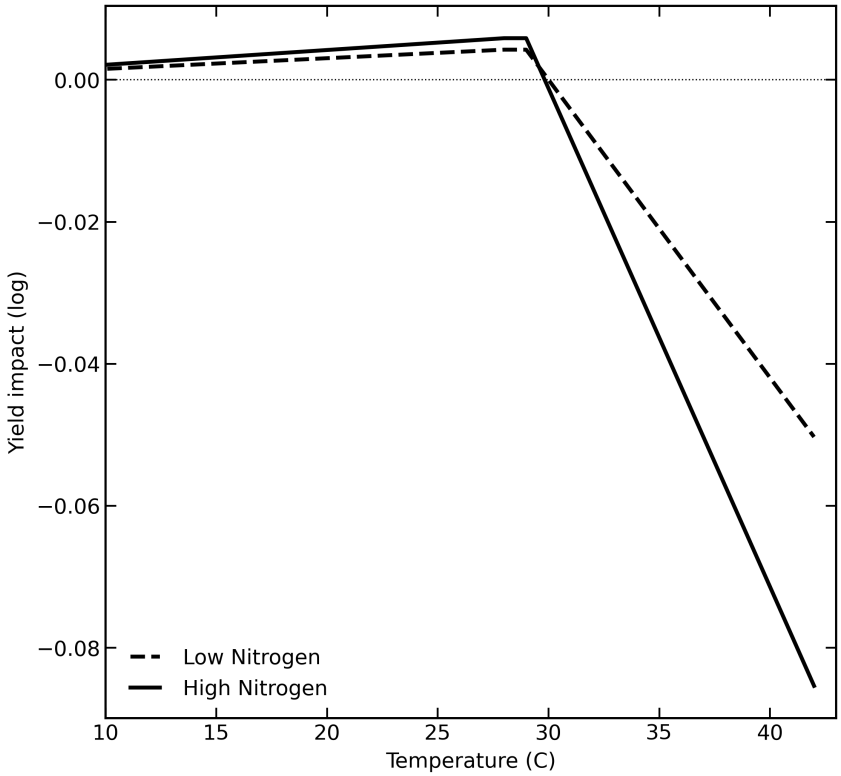


Figure 5.3: Panel model estimation of high temperature impacts on full-nitrogen and nitrogen-limited crop yields for maize in the North American Midwest. Following [180], the impact of a day at a certain temperature is estimated using a panel model with HDDs (above 29C) and GDDs and with county and year fixed effects.

between yields at elevated and present [CO_2] and at present and optimum nitrogen application levels and assign a percentage of production boost that is not achieved at current nitrogen application levels (see methods). We find a close association between nitrogen intensity and missed [CO_2] boost at the country and regional level. Regions with low nitrogen application levels are only able to capture a small fraction of the [CO_2] direct fertilization boost for all crops (Figure 5.5, blue curves) because the nitrogen stress is the limiting factor on crop growth.

The six crop models in this study (five for rice) agree in most cases on the nature of the interaction between weather and fertilizer application rates. The models differ quantitatively both in terms of nitrogen sensitivity under the historical conditions and the strength of the decline under warmer climatologies. The models disagree considerably more for rice at high

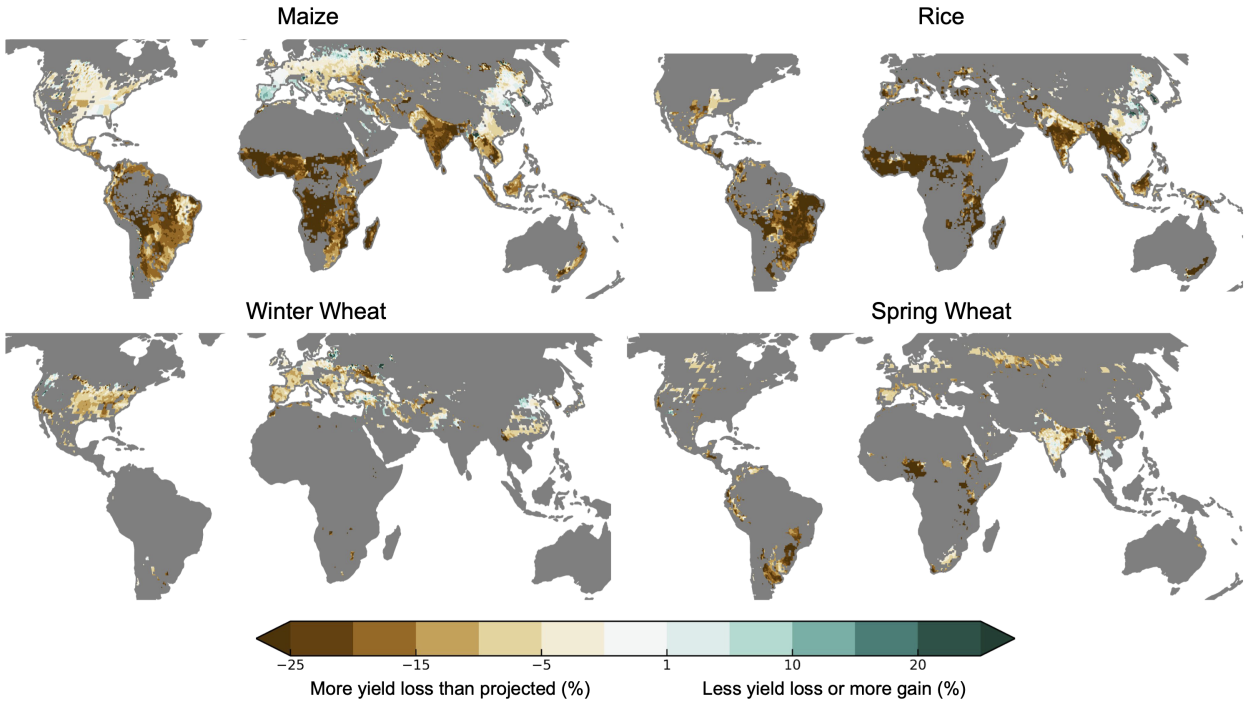


Figure 5.4: Potential mis-estimation of climate impacts on crops in nutrient limited areas. Map shows the difference between the temperature impact on each crop (across areas with current) cultivation in terms of mean percentage points between the full nitrogen (200 kg ha-1) and current nitrogen applications at 6 degrees of warming. Brown indicates areas which may experience more losses (or less gain) due to warming than may be expected if nitrogen levels continue to increase in the future from their current positions.

nitrogen application levels than for maize or wheat (Figure 5.2, middle panel). However, all the models included in this study show a steepening of yield damages under climate change for grains with higher nitrogen application levels.

Finally, we calculate future global production trajectories under high and low nitrogen applications and with and without accounting for [CO_2] impacts. A high nitrogen world stands to lose relatively more to temperature increases and gain relatively more to [CO_2] increases (Figure 5.6. Resultant impacts are therefore a product of two competing uncertainties with much spatial heterogeneity in which side of the tug-of-war prevails. Curiously, the current practice distribution of nitrogen looks much like high nitrogen pathway –likely due to the out sized production shares of those current high nitrogen regions.

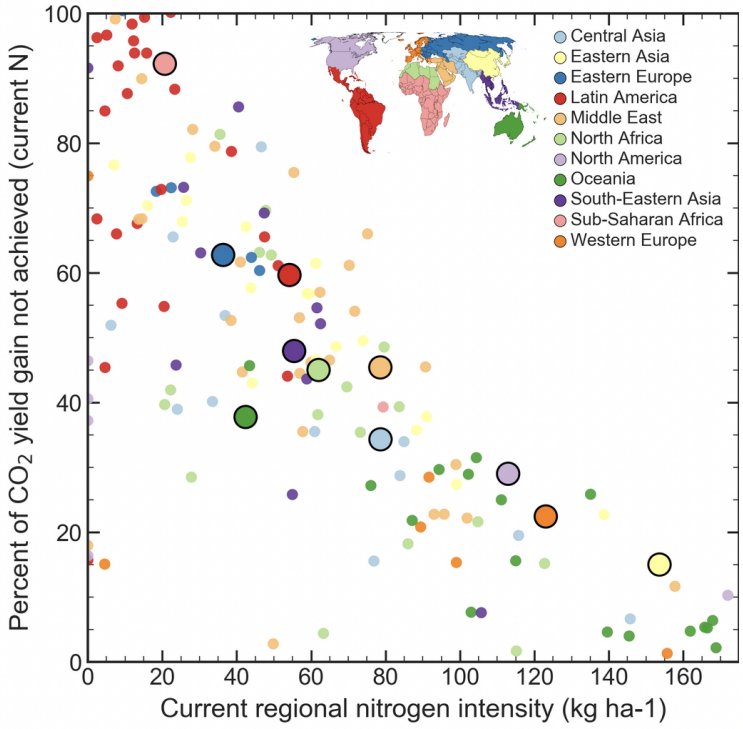


Figure 5.5: Current nitrogen application rates and potential missed [CO_2] benefit. Simulated production estimates for each country (small dots) and for the regional average (black outline dots) under end of century [CO_2 at both current nitrogen and 200 kg ha-1 nitrogen applications. Those regions with low nitrogen applications now are only able to capture a small portion of the potential yield gain from [CO_2] direct fertilization.

5.4 Discussion and Conclusions

Most crops in the extra tropics are presently grown in conditions near the optimal growing season temperature (in a median year) and therefore experience large yield gains from nitrogen fertilizer application. Nitrogen fertilizer applications vary widely in the real world. Crops that are currently grown near the optimal temperature in the median year will experience yield loss under climate change as they shift down the right side of the yield curve at a slope that is set by local nitrogen application levels, with steeper declines at higher fertilizer application levels. Some crops are presently grown in conditions below optimal growing season temperatures (e.g. less than 1400 GDD for maize) and will benefit from climate change as warming shifts those crops closer to optimum temperatures. Similarly in

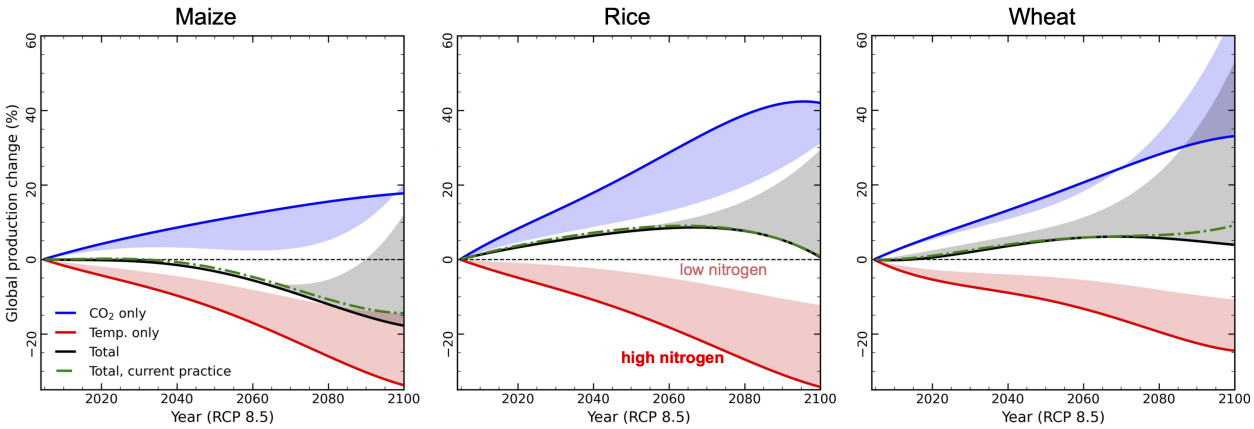


Figure 5.6: Global grain production change under possible high-end climate change for different levels of nitrogen applications. Wheat production aggregated at the caloric level globally over currently cultivated land. Lines indicate the percent change in calories under uniform 200 kg ha-1. Opposing side of the colored ribbons represents the change under 10 kg ha-1. Green dash represents impact under current nitrogen application rates held fixed into the future. All projections are mean projection across 5 models.

the positive case, the magnitude of the yield increase due to more favorable weather will be set in part by local nitrogen fertilizer application rates.

Two separate mechanisms are hypothesised to be leading nitrogen to modifying yield damages under warmer climatologies in addition to the stress hierarchy previously discussed. First, plants take up less nitrogen because they demand less and because there is less moisture in the soil. Second, plants transport less stored nitrogen from the leaves to the seeds during the grain filling portion of maturity. The temperature response in the latter case is likely more important during the grain filling stage, where reallocation of nitrogen from the leaves to the seeds can be inhibited by higher temperatures [85] due to a downregulation of the enzymes responsible for nitrogen transport and metabolism [58]. For example, temperatures above 22 C have decreased grain filling in wheat and have also been documented in field trials to reduce nitrogen assimilation [183]. Estimates for nitrogen transfer from the leaves to the seeds during grain filling illustrate that plants can only transfer as much as 60% [28] of the nitrogen stored in the leaves therefore changes in the growing season temperature distribution within the grain filling stage of maturity will be of the most importance.

These two mechanisms result in a decrease in yields and nitrogen uptake in most locations under warmer climatologies. Unless farmers adapt fertilizer application practices, climate change may lead to increased leaching and NH3 and NO_x emissions of which the respective negative effects on ocean eutrophication and on respiratory health outcomes are well documented [e.g. 53, 52]. We only track nitrogen losses within the growing season due to computational-cost restrictions and therefore cannot simulate nitrogen runoff and emission results since the additional nitrogen loss through emissions and leaching occurs outside of the growing season. Researchers need to modify the model framework to capture the full nitrogen cycle across the entire year in future studies.

The plant must balance nitrogen uptake and [CO_2] assimilation rates if the protein content of the grain is to remain constant under climate change. If plants take up and assimilate nitrogen at a rate that lags behind the yield change then the overall protein content of the grains may be reduced by climate change. This reduction in grain protein content has already been documented in FACE experiments at the field level for plants under high [CO_2] [e.g. 19]. Additional research is required to determine the exact nature of this response as we do not track grain protein content in many of the simulations.

Some important limitations with this study include holding current cultivated area constant, crop selection fixed, and precipitation at historical levels. We chose to hold the current cultivated area fixed in order to isolate the potential for fertilizer intensification. This assumption will degrade projections in high-latitude regions the most. We chose to focus on the temperature response in this study as opposed to the precipitation response for simplicity and because the latter is more uncertain in future climate projections. The simulations show the temperature effect dominates the water effect in regards to strength of the interaction with nitrogen. Irrigated crops (those with no water stress) show a similar though somewhat subdued response to their rain-fed counterparts. This is partly because irrigated crops are typically grown at higher temperatures, so progressively warmer temperatures brought

about by climate change cancel the benefits obtained from the additional water. Increased [CO_2] confounds the picture further because it modifies the way the plant uses water in an non-uniform way [32].

We show in this study that nitrogen availability strongly determines the magnitude of the impact climate change will have on grain production globally. Using a suite of process-based crop models, we show a near doubling of climate driven yield losses for each degree of warming under high nitrogen compared to low nitrogen application levels. Furthermore, unabated climate change may make it impossible to double production by nitrogen fertilizer intensification alone in some developing regions including Sub-Saharan Africa and South-East Asia. While nitrogen intensification has played the largest role in the globally tripling of grain production in the last 50 years, climate change may force some developing regions to focus on additional measures to meet production targets including expanding cultivation area, expanding irrigation, pesticide intensification, and cultivar switching. Further research should target testing nitrogen uptake and assimilation at the field level under different climate conditions so this relationship can be better constrained.

5.5 Data Availability

All data used in this analysis are publicly available. Simulation outputs from GGCMI phase 2 are available on at *https://zenodo.org/*. CMIP5 climate model outputs are publicly available at ISIMIP *https://data.isimip.org/*. AgMERRA data is available at reference: [171]. FAO data is available at reference: [46]. Nitrogen application data is available from reference: [83].

CHAPTER 6

CONCLUSIONS

In this work I presented the protocol and simulation outputs for a first-of-its-kind global crop model intercomparison project, the so-called Global Gridded Crop Model Intercomparison Project (GGCMI) Phase 2. This project gathered simulations under a harmonized protocol and ran simulations over a wide variety of climate and management inputs for major grains for the globe. I then presented the development and testing of a set of statistical emulators of these models and presented several use cases for said emulators. Next, I presented the first work to use an ensemble of process-based crop models to test how regions of peak yield may shift under expected warming. Finally, I show how applied nitrogen fertilizers modify the key climate response functions, namely that of temperature and [CO_2].

We conclude that the impacts of climate change are not equally distributed around the globe. The tropics stand to lose almost uniformly across all major crops even for moderate warming. Additionally, current high output agricultural regions in the extra-tropics may become less optimum under expected warming. These so-called 'breadbaskets' may shift poleward given changes grown season practices and genetics –both of which remain uncertain. Applied nitrogen fertilizers modify plant climate interactions. Regions of low inputs now may seem less sensitive to climate now than they will be in a more developed future.

While crop models are undoubtedly powerfully tools, care should be taken because their limitations are substantial. On the climate model side, part of the dynamics of fluid flow is resolved by solving the 3D Naiver-Stokes equations and part remains 'parameterized', that is, statistically approximated. In dynamic vegetation models however, no such analogous equation exists for photosynthesis and so all aspects of plant growth are in a sense, approximated. For example, soil representation varies greatly by model but generally contain nothing on microbes and have necessarily simple moisture dynamics. For this reason (and others) crop models show almost no sensitivity to excess water [105]. Evapotranspiration

(a key plant+soil process) parametrizations are based on one of two studies –one from 1965 and a updated version from 1972 [149] where meteorological data is good, but are sometimes based only on temperature. Thus we find crop model uncertainty outweighs climate model uncertainty for the future [128]. Crop models therefore need better links to experiments and observations [e.g. 168].

Nevertheless, several lines of analysis point to the necessity of the continued use of global process-based models to study the impacts of climate change including: [CO_2] impacts (which cannot be captured in empirical models), and the difference between the climatological and year-to-year temperatures responses, which may conflate temperature and water stress (not shown here). Crucially, crops grown under different production systems will likely respond to climate change quite differently. As shown here, the amount of applied fertilizers (which varies dramatically around the globe) modifies the plant response to climate change. Finally, changing genetics will play a very large and uncertain role [e.g. 216] which must be improved in models. Future work so continue model improvement efforts and develop ways to integrate remote and in situ observations directly into modeling frameworks.

REFERENCES

[1] Komlavi Akpoti, Amos T. Kabo-bah, and Sander J. Zwart. Agricultural land suitability analysis: State-of-the-art and outlooks for integration of climate change analysis. *Agricultural Systems*, 173:172–208, July 2019.

[2] W. B. Anderson, R. Seager, W. Baethgen, M. Cane, and L. You. Synchronous crop failures and climate-forced production variability. *Science Advances*, 5(7):eaaw1976, July 2019.

[3] Carlos Angulo, Reimund Rötter, Reiner Lock, Andreas Enders, Stefan Fronzek, and Frank Ewert. Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe. *Agricultural and Forest Meteorology*, 170:32–46, March 2013.

[4] S. Asseng, F. Ewert, P. Martre, R. P. Rötter, D. B. Lobell, D. Cammarano, B. A. Kimball, M. J. Ottman, G. W. Wall, J. W. White, M. P. Reynolds, P. D. Alderman, P. V. V. Prasad, P. K. Aggarwal, J. Anothai, B. Basso, C. Biernath, A. J. Challinor, G. De Sanctis, J. Doltra, E. Fereres, M. Garcia-Vila, S. Gayler, G. Hoogenboom, L. A. Hunt, R. C. Izaurralde, M. Jabloun, C. D. Jones, K. C. Kersebaum, A-K. Koehler, C. Müller, S. Naresh Kumar, C. Nendel, G. O'Leary, J. E. Olesen, T. Palosuo, E. Priesack, E. Eyshi Rezaei, A. C. Ruane, M. A. Semenov, I. Shcherbak, C. Stöckle, P. Stratonovitch, T. Streck, I. Supit, F. Tao, P. J. Thorburn, K. Waha, E. Wang, D. Wallach, J. Wolf, Z. Zhao, and Y. Zhu. Rising temperatures reduce global wheat production. *Nature Climate Change*, 5(2):143–147, February 2015.

[5] S. Asseng, F. Ewert, C. Rosenzweig, J. W. Jones, J. L. Hatfield, A. C. Ruane, K. J. Boote, P. J. Thorburn, R. P. Rötter, D. Cammarano, N. Brisson, B. Basso, P. Martre, P. K. Aggarwal, C. Angulo, P. Bertuzzi, C. Biernath, A. J. Challinor, J. Doltra, S. Gayler, R. Goldberg, R. Grant, L. Heng, J. Hooker, L. A. Hunt, J. Ingwersen, R. C. Izaurralde, K. C. Kersebaum, C. Müller, S. Naresh Kumar, C. Nendel, G. O'Leary, J. E. Olesen, T. M. Osborne, T. Palosuo, E. Priesack, D. Ripoche, M. A. Semenov, I. Shcherbak, P. Steduto, C. Stöckle, P. Stratonovitch, T. Streck, I. Supit, F. Tao, M. Travasso, K. Waha, D. Wallach, J. W. White, J. R. Williams, and J. Wolf. Uncertainty in simulating wheat yields under climate change. *Nature Climate Change*, 3(9):827–832, September 2013.

[6] Senthold Asseng, Pierre Martre, Andrea Maiorano, Reimund P. Rötter, Garry J. O'Leary, Glenn J. Fitzgerald, Christine Girousse, Rosella Motzo, Francesco Giunta, M. Ali Babar, Matthew P. Reynolds, Ahmed M. S. Kheir, Peter J. Thorburn, Katharina Waha, Alex C. Ruane, Pramod K. Aggarwal, Mukhtar Ahmed, Juraj Balkovič, Bruno Basso, Christian Biernath, Marco Bindi, Davide Cammarano, Andrew J. Challinor, Giacomo De Sanctis, Benjamin Dumont, Ehsan Eyshi Rezaei, Elias Fereres, Roberto Ferrise, Margarita Garcia-Vila, Sebastian Gayler, Yujing Gao, Heidi Horan, Gerrit Hoogenboom, R. César Izaurralde, Mohamed Jabloun, Curtis D. Jones, Belay T.

Kassie, Kurt-Christian Kersebaum, Christian Klein, Ann-Kristin Koehler, Bing Liu, Sara Minoli, Manuel Montesino San Martin, Christoph Müller, Soora Naresh Kumar, Claas Nendel, Jørgen Eivind Olesen, Taru Palosuo, John R. Porter, Eckart Priesack, Dominique Ripoche, Mikhail A. Semenov, Claudio Stöckle, Pierre Stratonovitch, Thilo Streck, Iwan Supit, Fulu Tao, Marijn Van der Velde, Daniel Wallach, Enli Wang, Heidi Webber, Joost Wolf, Liujun Xiao, Zhao Zhang, Zhigan Zhao, Yan Zhu, and Frank Ewert. Climate change impact and adaptation for wheat protein. *Global Change Biology*, 25(1):155–173, January 2019.

[7] Milkha S. Aulakh and Sukhdev S. Malhi. Interactions of Nitrogen with Other Nutrients and Water: Effect on Crop Yield and Quality, Nutrient Use Efficiency, Carbon Sequestration, and Environmental Pollution. volume 86 of *Advances in Agronomy*, pages 341–409. Academic Press, 2005. ISSN: 0065-2113.

[8] Justin E Bagley, Ankur R Desai, Paul A Dirmeyer, and Jonathan A Foley. Effects of land cover change on moisture availability and potential crop yield in the world's breadbaskets. *Environmental Research Letters*, 7(1):014009, March 2012.

[9] A. Bakhsh, R.S. Kanwar, and R.W. Malone. Role of landscape and hydrologic attributes in developing and interpreting yield clusters. *Geoderma*, 140(3):235–246, July 2007.

[10] Juraj Balkovič, Marijn van der Velde, Rastislav Skalský, Wei Xiong, Christian Folberth, Nikolay Khabarov, Alexey Smirnov, Nathaniel D. Mueller, and Michael Obersteiner. Global wheat production potentials and management flexibility under the representative concentration pathways. *Global and Planetary Change*, 122:107–121, November 2014.

[11] B. Basso, B. Dumont, B. Maestrini, I. Shcherbak, G. P. Robertson, J. R. Porter, P. Smith, K. Paustian, P. R. Grace, S. Asseng, S. Bassu, C. Biernath, K. J. Boote, D. Cammarano, G. De Sanctis, J.-L. Durand, F. Ewert, S. Gayler, D. W. Hyndman, J. Kent, P. Martre, C. Nendel, E. Priesack, D. Ripoche, A. C. Ruane, J. Sharp, P. J. Thorburn, J. L. Hatfield, J. W. Jones, and C. Rosenzweig. Soil Organic Carbon and Nitrogen Feedbacks on Crop Yields under Climate Change. *ael*, 3(1):0, 2018.

[12] N.H. Batjes. Harmonized soil property values for broad-scale modelling (wise30sec) with estimates of global soil carbon stocks. *Geoderma*, 269:61 – 68, 2016.

[13] Thomas Arthur Blair. A statistical study of weather factors affecting the yield of winter wheat in Ohio. *Monthly Weather Review*, 47(12):714–715, 1919.

[14] Élodie Blanc. Statistical emulators of maize, rice, soybean and wheat yields from global gridded crop models. *Agricultural and Forest Meteorology*, 236:145–161, April 2017.

[15] Elodie Blanc and Benjamin Sultan. Emulating maize yields from global gridded crop models using statistical estimates. *Agricultural and Forest Meteorology*, 214-215:134–147, December 2015.

[16] Benjamin Leon Bodirsky, Susanne Rolinski, Anne Biewald, Isabelle Weindl, Alexander Popp, and Hermann Lotze-Campen. Global Food Demand Scenarios for the 21st Century. *PLOS ONE*, 10(11):e0139201, November 2015.

[17] Elisabeth Bölenius, Bo Stenberg, and Johan Arvidsson. Within field cereal yield variability as affected by soil physical properties and weather variations – A case study in east central Sweden. *Geoderma Regional*, 11:96–103, 2017.

[18] Kenneth J. Boote, James W. Jones, Jeffrey W. White, Senthold Asseng, and Jon I. Lizaso. Putting mechanisms into crop production models: Putting mechanisms into crop production models. *Plant, Cell & Environment*, 36(9):1658–1672, September 2013.

[19] Peter Buchner, Michael Tausz, Rebecca Ford, Audrey Leo, Glenn J. Fitzgerald, Malcolm J. Hawkesford, and Sabine Tausz-Posch. Expression patterns of C- and N-metabolism related genes in wheat are changed during senescence under elevated CO2 in dry-land agriculture. *Plant Science*, 236:239–249, July 2015.

[20] Stefano Castruccio, David J. McInerney, Michael L. Stein, Feifei Liu Crouch, Robert L. Jacob, and Elisabeth J. Moyer. Statistical Emulation of Climate Model Projections Based on Precomputed GCM Runs*. *Journal of Climate*, 27(5):1829–1844, March 2014.

[21] Chatrina Caviezel, Matthias Hunziker, and Nikolaus Kuhn. Bequest of the Norseman—The Potential for Agricultural Intensification and Expansion in Southern Greenland under Climate Change. *Land*, 6(4):87, December 2017.

[22] A. J. Challinor, J. Watson, D. B. Lobell, S. M. Howden, D. R. Smith, and N. Chhetri. A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change*, 4(4):287–291, April 2014.

[23] A.J. Challinor, T.R. Wheeler, P.Q. Craufurd, J.M. Slingo, and D.I.F. Grimes. Design and optimisation of a large-area process-based model for annual crops. *Agricultural and Forest Meteorology*, 124(1-2):99–120, July 2004.

[24] Andrew J. Challinor, Christoph Müller, Senthold Asseng, Chetan Deva, Kathryn Jane Nicklin, Daniel Wallach, Eline Vanuytrecht, Stephen Whitfield, Julian Ramirez-Villegas, and Ann-Kristin Koehler. Improving the use of crop models for risk assessment and climate change adaptation. *Agricultural Systems*, 159:296–306, January 2018.

[25] Xavier Cirera and Edoardo Masset. Income distribution trends and future food demand. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1554):2821–2834, September 2010.

[26] Arnaud Costinot, Dave Donaldson, and Cory Smith. Evolving Comparative Advantage and the Impact of Climate Change in Agricultural Markets: Evidence from 1.7 Million Fields around the World. *journal of political economy*, page 44, 2016.

[27] C.T. de Wit. Transpiration and crop yields. *Verslagen van Landbouwkundige Onderzoekingen : 64.6*, (2), 1957.

[28] Jason DeBruin, Carlos D. Messina, Eugenia Munaro, Karen Thompson, Caitlin Conlon-Beckner, Lynne Fallis, David M. Sevenich, Rajeev Gupta, and Kanwarpal S. Dhugga. N distribution in maize plant as a marker for grain yield and limits on its remobilization after flowering. *Plant Breeding*, pages n/a–n/a, March 2013.

[29] Mark Decker, Shaoxiu Ma, and Andy Pitman. Local land–atmosphere feedbacks limit irrigation demand. *Environmental Research Letters*, 12(5):054003, May 2017.

[30] D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597, April 2011.

[31] R. DeFries and C. Rosenzweig. Toward a whole-landscape approach for sustainable land use in the tropics. *Proceedings of the National Academy of Sciences*, 107(46):19627–19632, November 2010.

[32] Delphine Deryng, Joshua Elliott, Christian Folberth, Christoph Müller, Thomas A. M. Pugh, Kenneth J. Boote, Declan Conway, Alex C. Ruane, Dieter Gerten, James W. Jones, Nikolay Khabarov, Stefan Olin, Sibyll Schaphoff, Erwin Schmid, Hong Yang, and Cynthia Rosenzweig. Regional disparities in the beneficial effects of rising CO2 concentrations on crop water productivity. *Nature Climate Change*, 6(8):786–790, August 2016.

[33] M Dury, A Hambuckers, P Warnant, A Henrot, E Favre, M Ouberdous, and L François. Responses of European forest ecosystems to 21st century climate: assessing changes in interannual variability and fire intensity. *iForest - Biogeosciences and Forestry*, 4(2):82–99, April 2011.

[34] J. Elliott, C. Müller, D. Deryng, J. Chryssanthacopoulos, K. J. Boote, M. Büchner, I. Foster, M. Glotter, J. Heinke, T. Iizumi, R. C. Izaurralde, N. D. Mueller, D. K. Ray, C. Rosenzweig, A. C. Ruane, and J. Sheffield. The Global Gridded Crop Model Intercomparison: data and modeling protocols for Phase 1 (v1.0). *Geoscientific Model Development*, 8(2):261–277, February 2015.

[35] Joshua Elliott, Delphine Deryng, Christoph Müller, Katja Frieler, Markus Konzmann, Dieter Gerten, Michael Glotter, Martina Flörke, Yoshihide Wada, Neil Best, Stephanie Eisner, Balázs M. Fekete, Christian Folberth, Ian Foster, Simon N. Gosling, Ingjerd

Haddeland, Nikolay Khabarov, Fulco Ludwig, Yoshimitsu Masaki, Stefan Olin, Cynthia Rosenzweig, Alex C. Ruane, Yusuke Satoh, Erwin Schmid, Tobias Stacke, Qiuhong Tang, and Dominik Wisser. Constraints and potentials of future irrigation water availability on agricultural production under climate change. *Proceedings of the National Academy of Sciences*, 111(9):3239–3244, March 2014.

[36] Joshua Elliott, David Kelly, James Chryssanthacopoulos, Michael Glotter, Kanika Jhunjhnuwala, Neil Best, Michael Wilde, and Ian Foster. The parallel system for integrating impact models and sectors (pSIMS). *Environmental Modelling & Software*, 62:509–516, December 2014.

[37] J. W. Erisman, M. A. Sutton, J. N. Galloway, and K. Zbigniew. How a century of ammonia synthesis changed the world. *nature geoscience*, 1:4, 2008.

[38] FAO/IIASA/ISRIC/ISSCAS/JRC. Harmonized world soil database (version 1.2), 2012.

[39] R. Ferrise, M. Moriondo, and M. Bindi. Probabilistic assessments of climate change impacts on durum wheat in the Mediterranean region. *Natural Hazards and Earth System Sciences*, 11(5):1293–1302, May 2011.

[40] G. Fischer, F. Nachtergaele, S. Prieler, H.T. van Velthuizen, L. Verelst, and D. Wiberg. Global Agro-ecological Zones Assessment for Agriculture (GAEZ), 2008.

[41] C. Folberth, A. Baklanov, J. Balkovič, R. Skalský, N. Khabarov, and M. Obersteiner. Spatio-temporal downscaling of gridded crop model yield estimates based on machine learning. *Agricultural and Forest Meteorology*, 264:1–15, January 2019.

[42] Christian Folberth, Thomas Gaiser, Karim C. Abbaspour, Rainer Schulin, and Hong Yang. Regionalization of a large-scale crop growth model for sub-Saharan Africa: Model setup, evaluation, and estimation of maize yields. *Agriculture, Ecosystems & Environment*, 151:21–33, April 2012.

[43] J. A. Foley. Global Consequences of Land Use. *Science*, 309(5734):570–574, July 2005.

[44] Jonathan A. Foley, Navin Ramankutty, Kate A. Brauman, Emily S. Cassidy, James S. Gerber, Matt Johnston, Nathaniel D. Mueller, Christine O'Connell, Deepak K. Ray, Paul C. West, Christian Balzer, Elena M. Bennett, Stephen R. Carpenter, Jason Hill, Chad Monfreda, Stephen Polasky, Johan Rockström, John Sheehan, Stefan Siebert, David Tilman, and David P. M. Zaks. Solutions for a cultivated planet. *Nature*, 478(7369):337–342, October 2011.

[45] Food and Agriculture Organization of the United Nations. How to feed the world in 2050. Technical report, FAO, 2009.

[46] Food and Agriculture Organization of the United Nations (FAO). FAOSTAT database, 2020.

[47] James A. Franke, Christoph Müller, Joshua Elliott, Alex C. Ruane, Jonas Jägermeyr, Juraj Balkovic, Philippe Ciais, Marie Dury, Pete D. Falloon, Christian Folberth, Louis François, Tobias Hank, Munir Hoffmann, R. Cesar Izaurralde, Ingrid Jacquemin, Curtis Jones, Nikolay Khabarov, Marian Koch, Michelle Li, Wenfeng Liu, Stefan Olin, Meridel Phillips, Thomas A. M. Pugh, Ashwan Reddy, Xuhui Wang, Karina Williams, Florian Zabel, and Elisabeth J. Moyer. The GGCMI Phase 2 experiment: global gridded crop model simulations under uniform changes in CO2, temperature, water, and nitrogen levels (protocol version 1.0). *Geoscientific Model Development*, 13(5):2315–2336, May 2020.

[48] James A. Franke, Christoph Müller, Joshua Elliott, Alex C. Ruane, Jonas Jägermeyr, Abigail Snyder, Marie Dury, Pete D. Falloon, Christian Folberth, Louis François, Tobias Hank, R. Cesar Izaurralde, Ingrid Jacquemin, Curtis Jones, Michelle Li, Wenfeng Liu, Stefan Olin, Meridel Phillips, Thomas A. M. Pugh, Ashwan Reddy, Karina Williams, Ziwei Wang, Florian Zabel, and Elisabeth J. Moyer. The GGCMI Phase 2 emulators: global gridded crop model responses to changes in CO2, temperature, water, and nitrogen (version 1.0). *Geoscientific Model Development*, 13(9):3995–4018, September 2020.

[49] James A. Franke, Christoph Müller, Sara Minoli, Joshua Elliott, Christian Folberth, Charles Gardner, Tobias Hank, Roberto Cesar Izaurralde, Jonas Jägermeyr, Curtis D. Jones, Wenfeng Liu, Stefan Olin, Thomas A.M. Pugh, Alex C. Ruane, Haynes Stephens, Florian Zabel, and Elisabeth J. Moyer. Agricultural breadbaskets shift poleward given adaptive farmer behavior under climate change. *Global Change Biology*, page gcb.15868, October 2021.

[50] Katja Frieler, Stefan Lange, Franziska Piontek, Christopher P. O. Reyer, Jacob Schewe, Lila Warszawski, Fang Zhao, Louise Chini, Sebastien Denvil, Kerry Emanuel, Tobias Geiger, Kate Halladay, George Hurtt, Matthias Mengel, Daisuke Murakami, Sebastian Ostberg, Alexander Popp, Riccardo Riva, Miodrag Stevanovic, Tatsuo Suzuki, Jan Volkholz, Eleanor Burke, Philippe Ciais, Kristie Ebi, Tyler D. Eddy, Joshua Elliott, Eric Galbraith, Simon N. Gosling, Fred Hattermann, Thomas Hickler, Jochen Hinkel, Christian Hof, Veronika Huber, Jonas Jägermeyr, Valentina Krysanova, Rafael Marcé, Hannes Müller Schmied, Ioanna Mouratiadou, Don Pierson, Derek P. Tittensor, Robert Vautard, Michelle van Vliet, Matthias F. Biber, Richard A. Betts, Benjamin Leon Bodirsky, Delphine Deryng, Steve Frolking, Chris D. Jones, Heike K. Lotze, Hermann Lotze-Campen, Ritvik Sahajpal, Kirsten Thonicke, Hanqin Tian, and Yoshiki Yamagata. Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). *Geoscientific Model Development*, 10(12):4321–4345, November 2017.

[51] Stefan Fronzek, Nina Pirttioja, Timothy R. Carter, Marco Bindi, Holger Hoffmann, Taru Palosuo, Margarita Ruiz-Ramos, Fulu Tao, Miroslav Trnka, Marco Acutis, Senthold Asseng, Piotr Baranowski, Bruno Basso, Per Bodin, Samuel Buis, Davide

Cammarano, Paola Deligios, Marie-France Destain, Benjamin Dumont, Frank Ewert, Roberto Ferrise, Louis François, Thomas Gaiser, Petr Hlavinka, Ingrid Jacquemin, Kurt Christian Kersebaum, Chris Kollas, Jaromir Krzyszczak, Ignacio J. Lorite, Julien Minet, M. Ines Minguez, Manuel Montesino, Marco Moriondo, Christoph Müller, Claas Nendel, Isik Öztürk, Alessia Perego, Alfredo Rodríguez, Alex C. Ruane, Françoise Ruget, Mattia Sanna, Mikhail A. Semenov, Cezary Slawinski, Pierre Stratonovitch, Iwan Supit, Katharina Waha, Enli Wang, Lianhai Wu, Zhigan Zhao, and Reimund P. Rötter. Classifying multi-model wheat yield impact response surfaces showing sensitivity to temperature and precipitation change. *Agricultural Systems*, 159:209–224, January 2018.

[52] J. N. Galloway, A. R. Townsend, J. W. Erisman, M. Bekunda, Z. Cai, J. R. Freney, L. A. Martinelli, S. P. Seitzinger, and M. A. Sutton. Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions. *Science*, 320(5878):889–892, May 2008.

[53] James N. Galloway, John D. Aber, Jan Willem Erisman, Sybil P. Seitzinger, Robert W. Howarth, Ellis B. Cowling, and B. Jack Cosby. The Nitrogen Cascade. *BioScience*, 53(4):341, 2003.

[54] Franziska Gaupp, Jim Hall, Stefan Hochrainer-Stigler, and Simon Dadson. Changing risks of simultaneous global breadbasket failure. *Nature Climate Change*, 10(1):54–57, January 2020.

[55] M. Glotter, J. Elliott, D. McInerney, N. Best, I. Foster, and E. J. Moyer. Evaluating the utility of dynamical downscaling in agricultural impacts projections. *Proceedings of the National Academy of Sciences*, 111(24):8776–8781, June 2014.

[56] Michael J. Glotter, Elisabeth J. Moyer, Alex C. Ruane, and Joshua W. Elliott. Evaluating the Sensitivity of Agricultural Model Performance to Different Climate Inputs*. *Journal of Applied Meteorology and Climatology*, 55(3):579–594, March 2016.

[57] H. Charles J. Godfray, John R. Beddington, Ian R. Crute, Lawernce Haddad, David Lawernce, James F. Muir, Jules Pretty, Sherman Robinson, Sandy M. Thomas, and Camilla Toulmin. Food Security: The Challenge of Feeding 9 Billion People. *Science*, 327:812–818, 2010.

[58] Parul Goel and Anil Kumar Singh. Abiotic Stresses Downregulate Key Genes Involved in Nitrogen Uptake and Assimilation in Brassica juncea L. *PLOS ONE*, 10(11):e0143645, November 2015.

[59] Tobias Hank, Heike Bach, and Wolfram Mauser. Using a Remote Sensing-Supported Hydro-Agroecological Model for Field-Scale Simulation of Heterogeneous Crop Growth and Yield: Application for Wheat in Central Europe. *Remote Sensing*, 7(4):3934–3965, April 2015.

[60] Lee Hannah, Patrick R. Roehrdanz, Krishna Bahadur K. C., Evan D. G. Fraser, Camila I. Donatti, Leonardo Saenz, Timothy Max Wright, Robert J. Hijmans, Mark Mulligan, Aaron Berg, and Arnout van Soesbergen. The environmental consequences of climate-driven agricultural frontiers. *PLOS ONE*, 15(2):e0228305, February 2020.

[61] J. Hansen, M. Sato, and R. Ruedy. Radiative forcing and climate response. *Journal of Geophysical Research: Atmospheres*, 102(D6):6831–6864, March 1997.

[62] J. W. Hansen and J. W. Jones. Scaling-up crop models for climate variability applications. *Agricultural Systems*, 65(1):43–72, 2000.

[63] I. Harris, P.D. Jones, T.J. Osborn, and D.H. Lister. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset: UPDATED HIGH-RESOLUTION GRIDS OF MONTHLY CLIMATIC OBSERVATIONS. *International Journal of Climatology*, 34(3):623–642, March 2014.

[64] Tomoko Hasegawa, Shinichiro Fujimori, Petr Havlík, Hugo Valin, Benjamin Leon Bodirsky, Jonathan C. Doelman, Thomas Fellmann, Page Kyle, Jason F. L. Koopman, Hermann Lotze-Campen, Daniel Mason-D'Croz, Yuki Ochi, Ignacio Pérez Domínguez, Elke Stehfest, Timothy B. Sulser, Andrzej Tabeau, Kiyoshi Takahashi, Jun'ya Takakura, Hans van Meijl, Willem-Jan van Zeist, Keith Wiebe, and Peter Witzke. Risk of increased food insecurity under stringent global climate change mitigation policy. *Nature Climate Change*, 8(8):699–703, August 2018.

[65] Matz A. Haugen, Michael L. Stein, Elisabeth J. Moyer, and Ryan L. Sriver. Estimating Changes in Temperature Distributions in a Large Ensemble of Climate Simulations Using Quantile Regression. *Journal of Climate*, 31(20):8573–8588, October 2018.

[66] QiJin He and GuangSheng Zhou. The climatic suitability for maize cultivation in China. *Chinese Science Bulletin*, 57(4):395–403, February 2012.

[67] Wentian He, J. Y. Yang, W. Zhou, C. F. Drury, X. M. Yang, W. D. Reynolds, H. Wang, P. He, and Z. T. Li. Sensitivity analysis of crop yields, soil water contents and nitrogen leaching to precipitation, management practices and soil hydraulic properties in semi-arid and humid regions of Canada using the DSSAT model. *Nutrient Cycling in Agroecosystems*, 106(2):201–215, November 2016.

[68] Earl O. Heady. An Econometric Investigation of the Technology of Agricultural Production Functions. *Econometrica*, 25(2):249–268, 1957.

[69] S. Hempel, K. Frieler, L. Warszawski, J. Schewe, and F. Piontek. A trend-preserving bias correction - the ISI-MIP approach. *Earth System Dynamics Discussions*, 4(1):49–92, January 2013.

[70] T.W. Hertel. Climate change, agricultural trade and global food security.the state of agricultural commodity markets (soco) 2018: Background paper 9. Technical report, FAO, 2018.

[71] A Hoffman, A Kemanian, and C Forest. The response of maize, sorghum, and soybean yield to growing-phase climate revealed with machine learning. *Environmental Research Letters*, 15(9):094013, August 2020.

[72] P. B. Holden, N. R. Edwards, P. H. Garthwaite, K. Fraedrich, F. Lunkeit, E. Kirk, M. Labriet, A. Kanudia, and F. Babonneau. PLASIM-ENTSem v1.0: a spatio-temporal emulator of future climate change for impacts assessment. *Geoscientific Model Development*, 7(1):433–451, February 2014.

[73] M. M. Holland and C. M. Bitz. Polar amplification of climate change in coupled models. *Climate Dynamics*, 21(3-4):221–232, September 2003.

[74] A Holzkämper, P Calanca, and J Fuhrer. Statistical crop models: predicting the effects of temperature and precipitation changes. *Climate Research*, 51(1):11–21, January 2012.

[75] Dean P. Holzworth, Neil I. Huth, Peter G. deVoil, Eric J. Zurcher, Neville I. Herrmann, Greg McLean, Karine Chenu, Erik J. van Oosterom, Val Snow, Chris Murphy, Andrew D. Moore, Hamish Brown, Jeremy P.M. Whish, Shaun Verrall, Justin Fainges, Lindsay W. Bell, Allan S. Peake, Perry L. Poulton, Zvi Hochman, Peter J. Thorburn, Donald S. Gaydon, Neal P. Dalgliesh, Daniel Rodriguez, Howard Cox, Scott Chapman, Alastair Doherty, Edmar Teixeira, Joanna Sharp, Rogerio Cichota, Iris Vogeler, Frank Y. Li, Enli Wang, Graeme L. Hammer, Michael J. Robertson, John P. Dimes, Anthony M. Whitbread, James Hunt, Harm van Rees, Tim McClelland, Peter S. Carberry, John N.G. Hargreaves, Neil MacLeod, Cam McDonald, Justin Harsdorf, Sara Wedgwood, and Brian A. Keating. APSIM – Evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software*, 62:327–350, December 2014.

[76] S. M. Howden and S. Crimp. Assessing dangerous climate change impacts on Australia's wheat industry. *Modelling and Simulation Society of Australia and New Zealand,*, pages 505–511, 2005.

[77] Solomon Hsiang, Robert Kopp, Amir Jina, James Rising, Michael Delgado, Shashank Mohan, D. J. Rasmussen, Robert Muir-Wood, Paul Wilson, Michael Oppenheimer, Kate Larsen, and Trevor Houser. Estimating economic damage from climate change in the United States. *Science*, 356(6345):1362–1369, June 2017.

[78] Florian Humpenöder, Alexander Popp, Benjamin Leon Bodirsky, Isabelle Weindl, Anne Biewald, Hermann Lotze-Campen, Jan Philipp Dietrich, David Klein, Ulrich Kreidenweis, Christoph Müller, Susanne Rolinski, and Miodrag Stevanovic. Large-scale bioenergy production: how to resolve sustainability trade-offs? *Environmental Research Letters*, 13(2):024011, February 2018.

[79] Toshichika Iizumi, Wonsik Kim, and Motoki Nishimori. Modeling the Global Sowing and Harvesting Windows of Major Crops Around the Year 2000. *Journal of Advances in Modeling Earth Systems*, 11(1):99–112, January 2019.

[80] Toshichika Iizumi, Motoki Nishimori, and Masayuki Yokozawa. Diagnostics of Climate Model Biases in Summer Temperature and Warm-Season Insolation for the Simulation of Regional Paddy Rice Yield in Japan. *Journal of Applied Meteorology and Climatology*, 49(4):574–591, April 2010.

[81] Toshichika Iizumi, Masayuki Yokozawa, Gen Sakurai, Maria Isabel Travasso, Vladimir Romanenkov, Pascal Oettli, Terry Newby, Yasushi Ishigooka, and Jun Furuya. Historical changes in global yields: major cereal and legume crops from 1982 to 2006. *Global Ecology and Biogeography*, 23(3):346–357, March 2014.

[82] Torsten Ingestad. Nitrogen and Plant Growth; Maximum Efficiency of Nitrogen Fertilizers. *AMBIO*, 6:146–151, 1977.

[83] International Fertilizer Association. IFASTAT database, 2019.

[84] International Food Policy Research Institute. Global Spatially-Disaggregated Crop Production Statistics Data for 2010 Version 2.0, 2019.

[85] S. Ito, T. Hara, Y. Kawanami, T. Watanabe, K. Thiraporn, N. Ohtake, K. Sueyoshi, T. Mitsui, T. Fukuyama, Y. Takahashi, T. Sato, A. Sato, and T. Ohyama. Carbon and Nitrogen Transport during Grain Filling in Rice Under High-temperature Conditions. *Journal of Agronomy and Crop Science*, 195(5):368–376, October 2009.

[86] R.C. Izaurralde, J.R. Williams, W.B. McGill, N.J. Rosenberg, and M.C. Quiroga Jakas. Simulating soil C dynamics with EPIC: Model description and testing against long-term data. *Ecological Modelling*, 192(3-4):362–384, February 2006.

[87] Jonas Jägermeyr and Katja Frieler. Spatial variations in crop growing seasons pivotal to reproduce global fluctuations in maize and wheat yields. *Science Advances*, 4(11):eaat4517, November 2018.

[88] Jonas Jägermeyr, Christoph Müller, Alex C. Ruane, Joshua Elliott, Juraj Balkovic, Oscar Castillo, Babacar Faye, Ian Foster, Christian Folberth, James A. Franke, Kathrin Fuchs, Jose R. Guarin, Jens Heinke, Gerrit Hoogenboom, Toshichika Iizumi, Atul K. Jain, David Kelly, Nikolay Khabarov, Stefan Lange, Tzu-Shun Lin, Wenfeng Liu, Oleksandr Mialyk, Sara Minoli, Elisabeth J. Moyer, Masashi Okada, Meridel Phillips, Cheryl Porter, Sam S. Rabin, Clemens Scheer, Julia M. Schneider, Joep F. Schyns, Rastislav Skalsky, Andrew Smerald, Tommaso Stella, Haynes Stephens, Heidi Webber, Florian Zabel, and Cynthia Rosenzweig. Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nature Food*, November 2021.

[89] Jonas Jägermeyr, Amandine Pastor, Hester Biemans, and Dieter Gerten. Reconciling irrigated food production with environmental flows for Sustainable Development Goals implementation. *Nature Communications*, 8(1):15900, August 2017.

[90] Shrikant S. Jagtap and James W. Jones. Adaptation and evaluation of the cropgro-soybean model to predict regional yield and production. *Agriculture, Ecosystems and Environment*, 93(1-3):73–85, 2002.

[91] Charlotte Janssens, Petr Havlík, Tamás Krisztin, Justin Baker, Stefan Frank, Tomoko Hasegawa, David Leclère, Sara Ohrel, Shaun Ragnauth, Erwin Schmid, Hugo Valin, Nicole Van Lipzig, and Miet Maertens. Global hunger and climate change adaptation through international trade. *Nature Climate Change*, 10(9):829–835, September 2020.

[92] C. D. Jones, J. K. Hughes, N. Bellouin, S. C. Hardiman, G. S. Jones, J. Knight, S. Liddicoat, F. M. O'Connor, R. J. Andres, C. Bell, K.-O. Boo, A. Bozzo, N. Butchart, P. Cadule, K. D. Corbin, M. Doutriaux-Boucher, P. Friedlingstein, J. Gornall, L. Gray, P. R. Halloran, G. Hurtt, W. J. Ingram, J.-F. Lamarque, R. M. Law, M. Meinshausen, S. Osprey, E. J. Palin, L. Parsons Chini, T. Raddatz, M. G. Sanderson, A. A. Sellar, A. Schurer, P. Valdes, N. Wood, S. Woodward, M. Yoshioka, and M. Zerroukat. The HadGEM2-ES implementation of CMIP5 centennial simulations. *Geoscientific Model Development*, 4(3):543–570, July 2011.

[93] James W. Jones, John M. Antle, Bruno Basso, Kenneth J. Boote, Richard T. Conant, Ian Foster, H. Charles J. Godfray, Mario Herrero, Richard E. Howitt, Sander Janssen, Brian A. Keating, Rafael Munoz-Carpena, Cheryl H. Porter, Cynthia Rosenzweig, and Tim R. Wheeler. Brief history of agricultural systems modeling. *Agricultural Systems*, 155:240–254, July 2017.

[94] J.W Jones, G Hoogenboom, C.H Porter, K.J Boote, W.D Batchelor, L.A Hunt, P.W Wilkens, U Singh, A.J Gijsman, and J.T Ritchie. The DSSAT cropping system model. *European Journal of Agronomy*, 18(3-4):235–265, January 2003.

[95] John Kearney. Food consumption trends and drivers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1554):2793–2807, September 2010.

[96] B.A Keating, P.S Carberry, G.L Hammer, M.E Probert, M.J Robertson, D Holzworth, N.I Huth, J.N.G Hargreaves, H Meinke, Z Hochman, G McLean, K Verburg, V Snow, J.P Dimes, M Silburn, E Wang, S Brown, K.L Bristow, S Asseng, S Chapman, R.L McCown, D.M Freebairn, and C.J Smith. An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy*, 18(3-4):267–288, January 2003.

[97] Myron King, Daniel Altdorff, Pengfei Li, Lakshman Galagedara, Joseph Holden, and Adrian Unc. Northward shift of the agricultural climate zone under 21st-century global climate change. *Scientific Reports*, 8(1):7904, December 2018.

[98] Patrycja Klusak, Matthew Agarwala, Matt Burke, Moritz Kraemer, and Kamiar Mohaddes. Rising temperatures, falling ratings: The effect of climate change on sovereign creditworthiness. *SSRN Electronic Journal*, 2021.

[99] Reto Knutti and Jan Sedláček. Robustness and uncertainties in the new CMIP5 climate model projections. *Nature Climate Change*, 3(4):369–373, April 2013.

[100] Benjamin Kipkemboi Kogo, Lalit Kumar, Richard Koech, and Champika S. Kariyawasam. Modelling Climate Suitability for Rainfed Maize Cultivation in Kenya

Using a Maximum Entropy (MaxENT) Approach. *Agronomy*, 9(11):727, November 2019.

[101] Kai Kornhuber, Dim Coumou, Elisabeth Vogel, Corey Lesk, Jonathan F. Donges, Jascha Lehmann, and Radley M. Horton. Amplified Rossby waves enhance risk of concurrent heatwaves in major breadbasket regions. *Nature Climate Change*, 10(1):48–53, January 2020.

[102] Peter Läderach, Julian Ramirez–Villegas, Carlos Navarro-Racines, Carlos Zelaya, Armando Martinez–Valle, and Andy Jarvis. Climate change adaptation of coffee production in space and time. *Climatic Change*, 141(1):47–62, March 2017.

[103] Luis Lassaletta, Gilles Billen, Bruna Grizzetti, Juliette Anglade, and Josette Garnier. 50 year trends in nitrogen use efficiency of world cropping systems: the relationship between yield and nitrogen input to cropland. *Environmental Research Letters*, 9(10):105011, October 2014.

[104] Luis Lassaletta, Gilles Billen, Bruna Grizzetti, Josette Garnier, Allison M. Leach, and James N. Galloway. Food and feed trade as a driver in the global nitrogen cycle: 50-year trends. *Biogeochemistry*, 118(1-3):225–241, April 2014.

[105] Yan Li, Kaiyu Guan, Gary D. Schnitkey, Evan DeLucia, and Bin Peng. Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. *Global Change Biology*, 25(7):2325–2337, 2019.

[106] Liebig, Justus. *Organic chemistry in its applications to agriculture and physiology.* Taylor and Walton, London, 1840.

[107] M. Lindeskog, A. Arneth, A. Bondeau, K. Waha, J. Seaquist, S. Olin, and B. Smith. Implications of accounting for land use in simulations of ecosystem carbon cycling in Africa. *Earth System Dynamics*, 4(2):385–407, November 2013.

[108] Bing Liu, Senthold Asseng, Christoph Müller, Frank Ewert, Joshua Elliott, David B. Lobell, Pierre Martre, Alex C. Ruane, Daniel Wallach, James W. Jones, Cynthia Rosenzweig, Pramod K. Aggarwal, Phillip D. Alderman, Jakarat Anothai, Bruno Basso, Christian Biernath, Davide Cammarano, Andy Challinor, Delphine Deryng, Giacomo De Sanctis, Jordi Doltra, Elias Fereres, Christian Folberth, Margarita Garcia-Vila, Sebastian Gayler, Gerrit Hoogenboom, Leslie A. Hunt, Roberto C. Izaurralde, Mohamed Jabloun, Curtis D. Jones, Kurt C. Kersebaum, Bruce A. Kimball, Ann-Kristin Koehler, Soora Naresh Kumar, Claas Nendel, Garry J. O'Leary, Jørgen E. Olesen, Michael J. Ottman, Taru Palosuo, P. V. Vara Prasad, Eckart Priesack, Thomas A. M. Pugh, Matthew Reynolds, Ehsan E. Rezaei, Reimund P. Rötter, Erwin Schmid, Mikhail A. Semenov, Iurii Shcherbak, Elke Stehfest, Claudio O. Stöckle, Pierre Stratonovitch, Thilo Streck, Iwan Supit, Fulu Tao, Peter Thorburn, Katharina Waha, Gerard W. Wall, Enli Wang, Jeffrey W. White, Joost Wolf, Zhigan Zhao, and Yan Zhu. Similar estimates of temperature impacts on global wheat yield by three independent methods. *Nature Climate Change*, 6(12):1130–1136, December 2016.

[109] Junguo Liu, Jimmy R. Williams, Alexander J.B. Zehnder, and Hong Yang. GEPIC – modelling wheat yield and crop water productivity with high resolution on a global scale. *Agricultural Systems*, 94(2):478–493, May 2007.

[110] Wenfeng Liu, Hong Yang, Christian Folberth, Xiuying Wang, Qunying Luo, and Rainer Schulin. Global investigation of impacts of PET methods on simulating crop-water relations for maize. *Agricultural and Forest Meteorology*, 221:164–175, May 2016.

[111] Wenfeng Liu, Hong Yang, Junguo Liu, Ligia B. Azevedo, Xiuying Wang, Zongxue Xu, Karim C. Abbaspour, and Rainer Schulin. Global assessment of nitrogen losses and trade-offs with yields from major crop cultivations. *Science of The Total Environment*, 572:526–537, December 2016.

[112] David B Lobell. Comparing estimates of climate change impacts from processbased and statistical crop models. *Environ. Res. Lett.*, page 13, 2017.

[113] David B Lobell and Marshall B Burke. Why are agricultural impacts of climate change so uncertain? The importance of temperature relative to precipitation. *Environmental Research Letters*, 3(3):034007, July 2008.

[114] David B. Lobell and Marshall B. Burke. On the use of statistical models to predict crop yield responses to climate change. *Agricultural and Forest Meteorology*, 150(11):1443–1452, October 2010.

[115] David B Lobell and Christopher B Field. Global scale climate–crop yield relationships and the impacts of recent warming. *Environmental Research Letters*, 2(1):014002, March 2007.

[116] David MacKay. Bayesian Interpolation. *Neural Computation*, 4(3):415–447, 1992.

[117] Wolfram Mauser, Gernot Klepper, Florian Zabel, Ruth Delzeit, Tobias Hank, Birgitta Putzenlechner, and Alvaro Calzadilla. Global biomass production potentials exceed expected future demand without the need for cropland expansion. *Nature Communications*, 6(1), December 2015.

[118] Sonali P. McDermid, Alexander C. Ruane, Cynthia Rosenzweig, Nicholas I. Hudson, Monica D. Morales, Prabodha Agalawatte, Shakeel Ahmad, L. R. Ahuja, Istiqlal Amien, Saseendran S. Anapalli, Jakarat Anothai, Senthold Asseng, Jody Biggs, Federico Bert, Patrick Bertuzzi, Virender S. Bhatia, Marco Bindi, Ian Broad, Davide Cammarano, Ramiro Carretero, Ashfaq Ahmad Chattha, Uran Chung, Stephanie Debats, Paola Deligios, Giacomo De Sanctis, Thanda Dhliwayo, Benjamin Dumont, Lyndon Estes, Frank Ewert, Roberto Ferrise, Thomas Gaiser, Guillermo Garcia, Sika Gbegbelegbe, Vellingiri Geethalakshmi, Edward Gerardeaux, Richard Goldberg, Brian Grant, Edgardo Guevara, Jonathan Hickman, Holger Hoffmann, Huanping Huang, Jamshad Hussain, Flavio Barbosa Justino, Asha S. Karunaratne, Ann-Kristin Koehler, Patrice K. Kouakou, Soora Naresh Kumar, Arunachalam Lakshmanan, Mark Liefering, Xiaomao Lin, Qunying Luo, Graciela Magrin, Marco Mancini, Fabio Ricardo

Marin, Anna Dalla Marta, Yuji Masutomi, Theodoros Mavromatis, Greg McLean, Santiago Meira, Monoranjan Mohanty, Marco Moriondo, Wajid Nasim, Lamyaa Negm, Francesca Orlando, Simone Orlandini, Isik Ozturk, Helena Maria Soares Pinto, Guillermo Podesta, Zhiming Qi, Johanna Ramarohetra, Muhammad Habib ur Rahman, Helene Raynal, Gabriel Rodriguez, Reimund Rötter, Vaishali Sharda, Lu Shuo, Ward Smith, Val Snow, Afshin Soltani, K. Srinivas, Benjamin Sultan, Dillip Kumar Swain, Fulu Tao, Kindie Tesfaye, Maria I. Travasso, Giacomo Trombi, Alex Topaj, Eline Vanuytrecht, Federico E. Viscarra, Syed Aftab Wajid, Enli Wang, Hong Wang, Jing Wang, Erandika Wijekoon, Lee Byun-Woo, Yang Xiaoguang, Ban Ho Young, Jin I. Yun, Zhigan Zhao, and Lareef Zubair. The AgMIP Coordinated Climate-Crop Modeling Project (C3MP): Methods and Protocols. In *Handbook of Climate Change and Agroecosystems*, pages 191–220. 2015.

[119] Justin M McGrath and David B Lobell. Regional disparities in the CO2 fertilization effect and implications for crop yields. *Environmental Research Letters*, 8(1):014054, March 2013.

[120] Sara Minoli, Dennis B. Egli, Susanne Rolinski, and Christoph Müller. Modelling cropping periods of grain crops at the global scale. *Global and Planetary Change*, 174:35–46, March 2019.

[121] Sara Minoli, Christoph Müller, Joshua Elliott, Alex C. Ruane, Jonas Jägermeyr, Florian Zabel, Marie Dury, Christian Folberth, Louis François, Tobias Hank, Ingrid Jacquemin, Wenfeng Liu, Stefan Olin, and Thomas A. M. Pugh. Global Response Patterns of Major Rainfed Crops to Adaptation by Maintaining Current Growing Periods and Irrigation. *Earth's Future*, 7(12):1464–1480, December 2019.

[122] Malcolm N Mistry, Ian Sue Wing, and Enrica De Cian. Simulated vs. empirical weather responsiveness of crop yields: US evidence and implications for the agricultural impacts of climate change. *Environmental Research Letters*, 12(7):075007, July 2017.

[123] Frances C. Moore, Uris Baldos, Thomas Hertel, and Delavane Diaz. New science of climate change impacts on agriculture implies higher social cost of carbon. *Nature Communications*, 8(1), December 2017.

[124] Frances C Moore, Uris Lantz C Baldos, and Thomas Hertel. Economic impacts of climate change on agriculture: a comparison of process-based and statistical yield models. *Environmental Research Letters*, 12(6):065008, June 2017.

[125] Nathaniel D. Mueller, James S. Gerber, Matt Johnston, Deepak K. Ray, Navin Ramankutty, and Jonathan A. Foley. Closing yield gaps through nutrient and water management. *Nature*, 490(7419):254–257, October 2012.

[126] Christoph Müller, Joshua Elliott, James Chryssanthacopoulos, Almut Arneth, Juraj Balkovic, Philippe Ciais, Delphine Deryng, Christian Folberth, Michael Glotter, Steven Hoek, Toshichika Iizumi, Roberto C Izaurralde, Curtis Jones, Nikolay Khabarov, Peter

Lawrence, Wenfeng Liu, Stefan Olin, Thomas A M Pugh, Deepak K Ray, Ashwan Reddy, Cynthia Rosenzweig, Alex C Ruane, Gen Sakurai, Erwin Schmid, Rastislav Skalsky, Carol X Song, Xuhui Wang, Allard de Wit, and Hong Yang. Global gridded crop model evaluation: benchmarking, skills, deficiencies and implications. *Geosci. Model Dev.*, page 20, 2017.

[127] Christoph Müller, Joshua Elliott, James Chryssanthacopoulos, Delphine Deryng, Christian Folberth, Thomas A M Pugh, and Erwin Schmid. Implications of climate mitigation for future agricultural production. *Environmental Research Letters*, 10(12):125004, December 2015.

[128] Christoph Müller, James Franke, Jonas Jägermeyr, Alex C Ruane, Joshua Elliott, Elisabeth Moyer, Jens Heinke, Pete D Falloon, Christian Folberth, Louis Francois, Tobias Hank, R César Izaurralde, Ingrid Jacquemin, Wenfeng Liu, Stefan Olin, Thomas A M Pugh, Karina Williams, and Florian Zabel. Exploring uncertainties in global crop yield projections in a large ensemble of crop models and CMIP5 and CMIP6 climate scenarios. *Environmental Research Letters*, 16(3):034040, March 2021.

[129] Christoph Müller and Richard D. Robertson. Projecting future crop productivity for global economic modeling. *Agricultural Economics*, 45(1):37–50, January 2014.

[130] Takuji Nakamura, Mitsuru Osaki, Takayoshi Koike, Yuko T. Hanba, Eitaro Wada, and Toshiaki Tadano. Effect of CO $_2$ enrichment on carbon and nitrogen interaction in wheat and soybean. *Soil Science and Plant Nutrition*, 43(4):789–798, December 1997.

[131] Gerald Nelson. *Food Security, Farming, and Climate Change to 2050: Scenarios, Results, Policy Options*. International Food Policy Research Institute, IFPRI Research Monographs, 2010.

[132] Gerald C. Nelson, Hugo Valin, Ronald D. Sands, Petr Havlík, Helal Ahammad, Delphine Deryng, Joshua Elliott, Shinichiro Fujimori, Tomoko Hasegawa, Edwina Heyhoe, Page Kyle, Martin Von Lampe, Hermann Lotze-Campen, Daniel Mason d'Croz, Hans van Meijl, Dominique van der Mensbrugghe, Christoph Müller, Alexander Popp, Richard Robertson, Sherman Robinson, Erwin Schmid, Christoph Schmitz, Andrzej Tabeau, and Dirk Willenbockel. Climate change effects on agriculture: Economic responses to biophysical shocks. *Proceedings of the National Academy of Sciences*, 111(9):3274–3279, March 2014.

[133] J.E. Newman. Climate Change Impacts on the Growing Season of the North American Corn Belt. *Biometeorology*, 7:128–142, 1980.

[134] A. O'Hagan. Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering & System Safety*, 91(10):1290–1300, 2006.

[135] S. Olin, G. Schurgers, M. Lindeskog, D. Wårlind, B. Smith, P. Bodin, J. Holmér, and A. Arneth. Modelling the response of yields and tissue C : N to changes in atmospheric

CO2; and N management in the main wheat regions of western Europe. *Biogeosciences*, 12(8):2489–2515, April 2015.

[136] Brian C. O'Neill, Elmar Kriegler, Kristie L. Ebi, Eric Kemp-Benedict, Keywan Riahi, Dale S. Rothman, Bas J. van Ruijven, Detlef P. van Vuuren, Joern Birkmann, Kasper Kok, Marc Levy, and William Solecki. The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environmental Change*, 42:169–180, January 2017.

[137] Ariel Ortiz-Bobea, Toby R. Ault, Carlos M. Carrillo, Robert G. Chambers, and David B. Lobell. Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change*, 11(4):306–312, April 2021.

[138] M. Osaki. Carbon-nitrogen interaction model in field crop production. *Plant and Soil*, 155/156:4, 1993.

[139] T. Osborne, J. Gornall, J. Hooker, K. Williams, A. Wiltshire, R. Betts, and T. Wheeler. JULES-crop: a parametrisation of crops in the Joint UK Land Environment Simulator. *Geoscientific Model Development*, 8(4):1139–1155, April 2015.

[140] Sebastian Ostberg, Jacob Schewe, Katelin Childers, and Katja Frieler. Changes in crop yields and their variability at different levels of global warming. *Earth System Dynamics*, 9(2):479–496, May 2018.

[141] Oluwole K Oyebamiji, Neil R Edwards, Philip B Holden, Paul H Garthwaite, Sibyll Schaphoff, and Dieter Gerten. Emulating global climate change impacts on crop yields. *Statistical Modelling: An International Journal*, 15(6):499–525, December 2015.

[142] Camille Parmesan. Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, 37(1):637–669, December 2006.

[143] N Pirttioja, Tr Carter, S Fronzek, M Bindi, H Hoffmann, T Palosuo, M Ruiz-Ramos, F Tao, M Trnka, M Acutis, S Asseng, P Baranowski, B Basso, P Bodin, S Buis, D Cammarano, P Deligios, Mf Destain, B Dumont, F Ewert, R Ferrise, L François, T Gaiser, P Hlavinka, I Jacquemin, Kc Kersebaum, C Kollas, J Krzyszczak, Ij Lorite, J Minet, Mi Minguez, M Montesino, M Moriondo, C Müller, C Nendel, I Öztürk, A Perego, A Rodríguez, Ac Ruane, F Ruget, M Sanna, Ma Semenov, C Slawinski, P Stratonovitch, I Supit, K Waha, E Wang, L Wu, Z Zhao, and Rp Rötter. Temperature and precipitation effects on wheat yield across a European transect: a crop model ensemble analysis using impact response surfaces. *Climate Research*, 65:87–105, September 2015.

[144] Andrew Poppick, Elisabeth J. Moyer, and Michael L. Stein. Estimating trends in the global mean temperature record. *arXiv:1607.03855 [stat]*, July 2016. arXiv: 1607.03855.

[145] Felix T. Portmann, Stefan Siebert, and Petra Döll. MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling: MONTHLY IRRIGATED AND RAINFED CROP AREAS. *Global Biogeochemical Cycles*, 24(1):n/a–n/a, March 2010.

[146] Vera Porwollik, Christoph Müller, Joshua Elliott, James Chryssanthacopoulos, Toshichika Iizumi, Deepak K. Ray, Alex C. Ruane, Almut Arneth, Juraj Balkovič, Philippe Ciais, Delphine Deryng, Christian Folberth, Roberto C. Izaurralde, Curtis D. Jones, Nikolay Khabarov, Peter J. Lawrence, Wenfeng Liu, Thomas A.M. Pugh, Ashwan Reddy, Gen Sakurai, Erwin Schmid, Xuhui Wang, Allard de Wit, and Xiuchen Wu. Spatial and temporal uncertainty of crop yield aggregations. *European Journal of Agronomy*, 88:10–21, August 2017.

[147] Sandra L. Postel, Gretchen C. Daily, and Paul R. Ehrlich. Human Appropriation of Renewable Fresh Water. *Science*, 271(5250):785–788, 1996. _eprint: https://www.science.org/doi/pdf/10.1126/science.271.5250.785.

[148] N. J. Potter, L. Zhang, P. C. D. Milly, T. A. McMahon, and A. J. Jakeman. Effects of rainfall seasonality and soil moisture capacity on mean annual water balance for Australian catchments: WATER BALANCE OF AUSTRALIAN CATCHMENTS. *Water Resources Research*, 41(6), June 2005.

[149] C. H. B. Priestley and Taylor, R. J. On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. *Monthly Weather Review*, page 12, 1972.

[150] T.A.M. Pugh, C. Müller, J. Elliott, D. Deryng, C. Folberth, S. Olin, E. Schmid, and A. Arneth. Climate analogues suggest limited potential for intensification of production on current croplands under climate change. *Nature Communications*, 7(1):12608, November 2016.

[151] Yue Qin, John T. Abatzoglou, Stefan Siebert, Laurie S. Huning, Amir AghaKouchak, Justin S. Mankin, Chaopeng Hong, Dan Tong, Steven J. Davis, and Nathaniel D. Mueller. Agricultural risks from changing snowmelt. *Nature Climate Change*, 10(5):459–465, May 2020.

[152] J. RäISäNEN and L. Ruokolainen. Probabilistic forecasts of near-term climate change based on a resampling ensemble technique. *Tellus A: Dynamic Meteorology and Oceanography*, 58(4):461–472, January 2006.

[153] Navin Ramankutty, Amato T. Evan, Chad Monfreda, and Jonathan A. Foley. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, 22(1), 2008.

[154] Nadiezhda Y. Z. Ramirez-Cabral, Lalit Kumar, and Farzin Shabani. Global alterations in areas of suitability for maize production from climate change and using a mechanistic species distribution model (CLIMEX). *Scientific Reports*, 7(1):5910, December 2017.

[155] Marco Ratto, Andrea Castelletti, and Andrea Pagano. Emulation techniques for the reduction and sensitivity analysis of complex environmental models. *Environmental Modelling & Software*, 34:1–4, June 2012.

[156] Deepak K. Ray, James S. Gerber, Graham K. MacDonald, and Paul C. West. Climate variation explains a third of global crop yield variability. *Nature Communications*, 6(1):5989, May 2015.

[157] Deepak K. Ray, Navin Ramankutty, Nathaniel D. Mueller, Paul C. West, and Jonathan A. Foley. Recent patterns of crop yield growth and stagnation. *Nature Communications*, 3(1):1293, January 2012.

[158] Saman Razavi, Bryan A. Tolson, and Donald H. Burn. Review of surrogate modeling in water resources: REVIEW. *Water Resources Research*, 48(7), July 2012.

[159] Keywan Riahi, Arnulf Grübler, and Nebojsa Nakicenovic. Scenarios of long-term socio-economic and environmental development under climate stabilization. *Technological Forecasting and Social Change*, 74(7):887–935, September 2007.

[160] Keywan Riahi, Shilpa Rao, Volker Krey, Cheolhung Cho, Vadim Chirkov, Guenther Fischer, Georg Kindermann, Nebojsa Nakicenovic, and Peter Rafaj. RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change*, 109(1-2):33–57, November 2011.

[161] Michael J Roberts, Noah O Braun, Thomas R Sinclair, David B Lobell, and Wolfram Schlenker. Comparing and combining process-based crop models and statistical models with some implications for climate change. *Environmental Research Letters*, 12(9):095010, September 2017.

[162] Johan Rockström, Will Steffen, Kevin Noone, Åsa Persson, F. Stuart Chapin, Eric F. Lambin, Timothy M. Lenton, Marten Scheffer, Carl Folke, Hans Joachim Schellnhuber, Björn Nykvist, Cynthia A. de Wit, Terry Hughes, Sander van der Leeuw, Henning Rodhe, Sverker Sörlin, Peter K. Snyder, Robert Costanza, Uno Svedin, Malin Falkenmark, Louise Karlberg, Robert W. Corell, Victoria J. Fabry, James Hansen, Brian Walker, Diana Liverman, Katherine Richardson, Paul Crutzen, and Jonathan A. Foley. A safe operating space for humanity. *Nature*, 461(7263):472–475, September 2009.

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

[164] Cynthia Rosenzweig. Potential CO2-induced climate effects on North American wheat-producing regions. *Climatic Change*, 7(4):367–389, December 1985.

[165] Cynthia Rosenzweig, Joshua Elliott, Delphine Deryng, Alex C. Ruane, Christoph Müller, Almut Arneth, Kenneth J. Boote, Christian Folberth, Michael Glotter, Nikolay Khabarov, Kathleen Neumann, Franziska Piontek, Thomas A. M. Pugh, Erwin Schmid, Elke Stehfest, Hong Yang, and James W. Jones. 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, March 2014.

[166] Cynthia Rosenzweig, Alex C. Ruane, John Antle, Joshua Elliott, Muhammad Ashfaq, Ashfaq Ahmad Chatta, Frank Ewert, Christian Folberth, Ibrahima Hathie, Petr Havlik, Gerrit Hoogenboom, Hermann Lotze-Campen, Dilys S. MacCarthy, Daniel Mason-D'Croz, Erik Mencos Contreras, Christoph Müller, Ignacio Perez-Dominguez, Meridel Phillips, Cheryl Porter, Rubi M. Raymundo, Ronald D. Sands, Carl-Friedrich Schleussner, Roberto O. Valdivia, Hugo Valin, and Keith Wiebe. Coordinating AgMIP data and models across global and regional scales for 1.5°C and 2.0°C assessments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2119):20160455, May 2018.

[167] Reimund P. Rötter, Timothy R. Carter, Jørgen E. Olesen, and John R. Porter. Crop–climate models need an overhaul. *Nature Climate Change*, 1(4):175–177, July 2011.

[168] R.P. Rötter, M. Appiah, E. Fichtler, K.C. Kersebaum, M. Trnka, and M.P Hoffmann. Linking modelling and experimentation to better capture crop impacts of agroclimatic extremes—A review. *Field Crops Research*, 221:142–156, May 2018.

[169] Ac Ruane, J Antle, J Elliott, C Folberth, G Hoogenboom, D Mason-D'Croz, C Müller, C Porter, Mm Phillips, Rm Raymundo, R Sands, Ro Valdivia, Jw White, K Wiebe, and C Rosenzweig. Biophysical and economic implications for agriculture of +1.5° and +2.0°C global warming using AgMIP Coordinated Global and Regional Assessments. *Climate Research*, 76(1):17–39, September 2018.

[170] Alex C. Ruane, L. DeWayne Cecil, Radley M. Horton, Román Gordón, Raymond McCollum, Douglas Brown, Brian Killough, Richard Goldberg, Adam P. Greeley, and Cynthia Rosenzweig. Climate change impact uncertainties for maize in Panama: Farm information, climate projections, and yield sensitivities. *Agricultural and Forest Meteorology*, 170:132–145, March 2013.

[171] Alex C. Ruane, Richard Goldberg, and James Chryssanthacopoulos. Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. *Agricultural and Forest Meteorology*, 200:233–248, January 2015.

[172] Alex C. Ruane, Nicholas I. Hudson, Senthold Asseng, Davide Camarrano, Frank Ewert, Pierre Martre, Kenneth J. Boote, Peter J. Thorburn, Pramod K. Aggarwal, Carlos Angulo, Bruno Basso, Patrick Bertuzzi, Christian Biernath, Nadine Brisson, Andrew J. Challinor, Jordi Doltra, Sebastian Gayler, Richard Goldberg, Robert F. Grant,

Lee Heng, Josh Hooker, Leslie A. Hunt, Joachim Ingwersen, Roberto C. Izaurralde, Kurt Christian Kersebaum, Soora Naresh Kumar, Christoph Müller, Claas Nendel, Garry O'Leary, Jørgen E. Olesen, Tom M. Osborne, Taru Palosuo, Eckart Priesack, Dominique Ripoche, Reimund P. Rötter, Mikhail A. Semenov, Iurii Shcherbak, Pasquale Steduto, Claudio O. Stöckle, Pierre Stratonovitch, Thilo Streck, Iwan Supit, Fulu Tao, Maria Travasso, Katharina Waha, Daniel Wallach, Jeffrey W. White, and Joost Wolf. Multi-wheat-model ensemble responses to interannual climate variability. *Environmental Modelling & Software*, 81:86–101, July 2016.

[173] Alex C. Ruane, Sonali McDermid, Cynthia Rosenzweig, Guillermo A. Baigorria, James W. Jones, Consuelo C. Romero, and L. DeWayne Cecil. Carbon-Temperature-Water change analysis for peanut production under climate change: a prototype for the AgMIP Coordinated Climate-Crop Modeling Project (C3MP). *Global Change Biology*, 20(2):394–407, February 2014.

[174] Alex C Ruane, Cynthia Rosenzweig, Senthold Asseng, Kenneth J Boote, Joshua Elliott, Frank Ewert, James W Jones, Pierre Martre, Sonali P McDermid, Christoph Müller, Abigail Snyder, and Peter J Thorburn. An AgMIP framework for improved agricultural representation in integrated assessment models. *Environmental Research Letters*, 12(12):125003, December 2017.

[175] Franz Rubel and Markus Kottek. Observed and projected climate shifts 1901-2100 depicted by world maps of the Köppen-Geiger climate classification. *Meteorologische Zeitschrift*, 19(2):135–141, April 2010.

[176] M. Ruiz-Ramos, R. Ferrise, A. Rodríguez, I. J. Lorite, M. Bindi, T. R. Carter, S. Fronzek, T. Palosuo, N. Pirttioja, P. Baranowski, S. Buis, D. Cammarano, Y. Chen, B. Dumont, F. Ewert, T. Gaiser, P. Hlavinka, H. Hoffmann, J. G. Höhn, F. Jurecka, K. C. Kersebaum, J. Krzyszczak, M. Lana, A. Mechiche-Alami, J. Minet, M. Montesino, C. Nendel, J. R. Porter, F. Ruget, M. A. Semenov, Z. Steinmetz, P. Stratonovitch, I. Supit, F. Tao, M. Trnka, A. de Wit, and R. P. Rötter. Adaptation response surfaces for managing wheat under perturbed climate and CO2 in a Mediterranean environment. *Agricultural Systems*, 159:260–274, 2018.

[177] William J. Sacks, Delphine Deryng, Jonathan A. Foley, and Navin Ramankutty. Crop planting dates: an analysis of global patterns: Global crop planting dates. *Global Ecology and Biogeography*, pages no–no, June 2010.

[178] Bernhard Schauberger, Sotirios Archontoulis, Almut Arneth, Juraj Balkovic, Philippe Ciais, Delphine Deryng, Joshua Elliott, Christian Folberth, Nikolay Khabarov, Christoph Müller, Thomas A. M. Pugh, Susanne Rolinski, Sibyll Schaphoff, Erwin Schmid, Xuhui Wang, Wolfram Schlenker, and Katja Frieler. Consistent negative response of US crops to high temperatures in observations and crop models. *Nature Communications*, 8:13931, January 2017.

[179] Jacob Schewe, Simon N. Gosling, Christopher Reyer, Fang Zhao, Philippe Ciais, Joshua Elliott, Louis Francois, Veronika Huber, Heike K. Lotze, Sonia I. Seneviratne, Michelle T. H. van Vliet, Robert Vautard, Yoshihide Wada, Lutz Breuer, Matthias Büchner, David A. Carozza, Jinfeng Chang, Marta Coll, Delphine Deryng, Allard de Wit, Tyler D. Eddy, Christian Folberth, Katja Frieler, Andrew D. Friend, Dieter Gerten, Lukas Gudmundsson, Naota Hanasaki, Akihiko Ito, Nikolay Khabarov, Hyungjun Kim, Peter Lawrence, Catherine Morfopoulos, Christoph Müller, Hannes Müller Schmied, René Orth, Sebastian Ostberg, Yadu Pokhrel, Thomas A. M. Pugh, Gen Sakurai, Yusuke Satoh, Erwin Schmid, Tobias Stacke, Jeroen Steenbeek, Jörg Steinkamp, Qiuhong Tang, Hanqin Tian, Derek P. Tittensor, Jan Volkholz, Xuhui Wang, and Lila Warszawski. State-of-the-art global models underestimate impacts from climate extremes. *Nature Communications*, 10(1), December 2019.

[180] Wolfram Schlenker and Michael J. Roberts. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106(37):15594–15598, September 2009.

[181] Wei Shangguan, Yongjiu Dai, Qingyun Duan, Baoyuan Liu, and Hua Yuan. A global soil data set for earth system modeling. *Journal of Advances in Modeling Earth Systems*, 6(1):249–263, March 2014.

[182] Thomas R. Sinclair and Thomas W. Rufty. Nitrogen and water resources commonly limit crop yield increases, not necessarily plant genetics. *Global Food Security*, 1(2):94–98, December 2012.

[183] Gw Singletary, R Banisadr, and Pl Keeling. Heat Stress During Grain Filling in Maize: Effects on Carbohydrate Storage and Metabolism. *Functional Plant Biology*, 21(6):829, 1994.

[184] Lindsey L. Sloat, Steven J. Davis, James S. Gerber, Frances C. Moore, Deepak K. Ray, Paul C. West, and Nathaniel D. Mueller. Climate adaptation by crop migration. *Nature Communications*, 11(1):1243, December 2020.

[185] Vaclav Smil. Nitrogen and Food Production: Proteins for Human Diets. 31(2):6, 2002.

[186] Vaclav Smil. Nitrogen cycle and world food production. *World Agriculutre*, 2:9–13, January 2011.

[187] B. Smith, D. Wårlind, A. Arneth, T. Hickler, P. Leadley, J. Siltberg, and S. Zaehle. Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences*, 11(7):2027–2054, April 2014.

[188] Abigail Snyder, Katherine V. Calvin, Meridel Phillips, and Alex C. Ruane. A crop yield change emulator for use in GCAM and similar models: Persephone v1.0. *Geoscientific Model Development*, 12(4):1319–1350, April 2019.

[189] Miodrag Stevanović, Alexander Popp, Hermann Lotze-Campen, Jan Philipp Dietrich, Christoph Müller, Markus Bonsch, Christoph Schmitz, Benjamin Leon Bodirsky, Florian Humpenöder, and Isabelle Weindl. The impact of high-end climate change on agricultural welfare. *Science Advances*, 2(8):e1501452, August 2016.

[190] W. M. Stewart, D. W. Dibb, A. E. Johnston, and T. J. Smyth. The Contribution of Commercial Fertilizer Nutrients to Food Production. *Agronomy Journal*, 97(1):1, 2005.

[191] Curtis B. Storlie, Laura P. Swiler, Jon C. Helton, and Cedric J. Sallaberry. Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliability Engineering & System Safety*, 94(11):1735–1763, 2009.

[192] Karl E. Taylor, Ronald J. Stouffer, and Gerald A. Meehl. An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, 93(4):485–498, April 2012.

[193] C. Tebaldi and D. B. Lobell. Towards probabilistic projections of climate change impacts on global crop yields. *Geophysical Research Letters*, 35(8):L08705, April 2008.

[194] Claudia Tebaldi, Kevin Debeire, Veronika Eyring, Erich Fischer, John Fyfe, Pierre Friedlingstein, Reto Knutti, Jason Lowe, Brian O'Neill, Benjamin Sanderson, Detlef van Vuuren, Keywan Riahi, Malte Meinshausen, Zebedee Nicholls, George Hurtt, Elmar Kriegler, Jean-Francois Lamarque, Gerald Meehl, Richard Moss, Susanne E. Bauer, Olivier Boucher, Victor Brovkin, Jean-Christophe Golaz, Silvio Gualdi, Huan Guo, Jasmin G. John, Slava Kharin, Tsuyoshi Koshiro, Libin Ma, Dirk Olivié, Swapna Panickal, Fangli Qiao, Nan Rosenbloom, Martin Schupfner, Roland Seferian, Zhenya Song, Christian Steger, Alistair Sellar, Neil Swart, Kaoru Tachiiri, Hiroaki Tatebe, Aurore Voldoire, Evgeny Volodin, Klaus Wyser, Xiaoge Xin, Rong Xinyao, Shuting Yang, Yongqiang Yu, and Tilo Ziehn. Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6. *Earth System Dynamics*, September 2020.

[195] The HadGEM2 Development Team: G. M. Martin, N. Bellouin, W. J. Collins, I. D. Culverwell, P. R. Halloran, S. C. Hardiman, T. J. Hinton, C. D. Jones, R. E. McDonald, A. J. McLaren, F. M. O'Connor, M. J. Roberts, J. M. Rodriguez, S. Woodward, M. J. Best, M. E. Brooks, A. R. Brown, N. Butchart, C. Dearden, S. H. Derbyshire, I. Dharssi, M. Doutriaux-Boucher, J. M. Edwards, P. D. Falloon, N. Gedney, L. J. Gray, H. T. Hewitt, M. Hobson, M. R. Huddleston, J. Hughes, S. Ineson, W. J. Ingram, P. M. James, T. C. Johns, C. E. Johnson, A. Jones, C. P. Jones, M. M. Joshi, A. B. Keen, S. Liddicoat, A. P. Lock, A. V. Maidens, J. C. Manners, S. F. Milton, J. G. L. Rae, J. K. Ridley, A. Sellar, C. A. Senior, I. J. Totterdell, A. Verhoef, P. L. Vidale, and A. Wiltshire. The HadGEM2 family of Met Office Unified Model climate configurations. *Geoscientific Model Development*, 4(3):723–757, September 2011.

[196] D. Tilman, C. Balzer, J. Hill, and B. L. Befort. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, 108(50):20260–20264, December 2011.

[197] Andrea Toreti, Delphine Deryng, Francesco N. Tubiello, Christoph Müller, Bruce A. Kimball, Gerald Moser, Kenneth Boote, Senthold Asseng, Thomas A. M. Pugh, Eline Vanuytrecht, Håkan Pleijel, Heidi Webber, Jean-Louis Durand, Frank Dentener, Andrej Ceglar, Xuhui Wang, Franz Badeck, Remi Lecerf, Gerard W. Wall, Maurits van den Berg, Petra Hoegy, Raul Lopez-Lozano, Matteo Zampieri, Stefano Galmarini, Garry J. O'Leary, Remy Manderscheid, Erik Mencos Contreras, and Cynthia Rosenzweig. Narrowing uncertainties in the effects of elevated CO2 on crops. *Nature Food*, 1(12):775–782, December 2020.

[198] Daniel Urban, Michael J. Roberts, Wolfram Schlenker, and David B. Lobell. Projected temperature changes indicate significant increase in interannual variability of U.S. maize yields: A Letter. *Climatic Change*, 112(2):525–533, May 2012.

[199] US Department of Agriculture. Summary Report: National Resources Inventory. Technical report, Natural Resources Conservation Service, Washington, DC, and Center for Survey Statistics and Methodology, Iowa State University, 2009.

[200] USDA National Agricultural Statistics Service. NASS quick stats, 2020.

[201] Vincent Vadez, Jens D. Berger, Tom Warkentin, Senthold Asseng, Pasala Ratnakumar, K. Poorna Chandra Rao, Pooran M. Gaur, Nathalie Munier-Jolain, Annabelle Larmure, Anne-Sophie Voisin, Hari C. Sharma, Suresh Pande, Mamta Sharma, Lakshman Krishnamurthy, and Mainassara Abdou Zaman. Adaptation of grain legumes to climate change: a review. *Agronomy for Sustainable Development*, 32(1):31–44, January 2012.

[202] X.P.C. Vergé, C. De Kimpe, and R.L. Desjardins. Agricultural production, greenhouse gas emissions and mitigation potential. *Agricultural and Forest Meteorology*, 142(2-4):255–269, February 2007.

[203] Werner von Bloh, Sibyll Schaphoff, Christoph Müller, Susanne Rolinski, Katharina Waha, and Sönke Zaehle. Implementing the nitrogen cycle into the dynamic global vegetation, hydrology, and crop growth model LPJmL (version 5.0). *Geoscientific Model Development*, 11(7):2789–2812, July 2018.

[204] K. Waha, L. G. J. van Bussel, C. Müller, and A. Bondeau. Climate-driven simulation of global crop sowing dates: Simulation of global sowing dates. *Global Ecology and Biogeography*, 21(2):247–259, February 2012.

[205] Sherrie Wang, Stefania Di tommaso, Jillian M. Deines, and David B. Lobel. Mapping twenty years of corn and soybean across the US Midwest using the Landsat archive. *Scientific Data*, 7(307):14, 2020.

[206] R. A. Warrick. Carbon Dioxide, Climatic Change and Agriculture. *The Geographical Journal*, 154(2):221, July 1988.

[207] Lila Warszawski, Katja Frieler, Veronika Huber, Franziska Piontek, Olivia Serdeczny, and Jacob Schewe. The Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP): Project framework. *Proceedings of the National Academy of Sciences*, 111(9):3228–3232, March 2014.

[208] Graham P. Weedon, Gianpaolo Balsamo, Nicolas Bellouin, Sandra Gomes, Martin J. Best, and Pedro Viterbo. The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, 50(9):7505–7514, September 2014.

[209] T. Wheeler and J. von Braun. Climate Change Impacts on Global Food Security. *Science*, 341(6145):508–513, August 2013.

[210] Keith Wiebe, Hermann Lotze-Campen, Ronald Sands, Andrzej Tabeau, Dominique van der Mensbrugghe, Anne Biewald, Benjamin Bodirsky, Shahnila Islam, Aikaterini Kavallari, Daniel Mason-D'Croz, Christoph Müller, Alexander Popp, Richard Robertson, Sherman Robinson, Hans van Meijl, and Dirk Willenbockel. Climate change impacts on agriculture in 2050 under a range of plausible socioeconomic and emissions scenarios. *Environmental Research Letters*, 10(8):085010, August 2015.

[211] K. E. Williams and P. D. Falloon. Sources of interannual yield variability in JULES-crop and implications for forcing with seasonal weather forecasts. *Geoscientific Model Development*, 8(12):3987–3997, December 2015.

[212] Karina Williams, Jemma Gornall, Anna Harper, Andy Wiltshire, Debbie Hemming, Tristan Quaife, Tim Arkebauer, and David Scoby. Evaluation of JULES-crop performance against site observations of irrigated maize from Mead, Nebraska. *Geosci. Model Dev.*, page 30, 2017.

[213] J Wolf and M.van Oijen. Modelling the dependence of European potato yields on changes in climate and CO2. *Agricultural and Forest Meteorology*, 112(3-4):217–231, October 2002.

[214] X. Wu, N. Vuichard, P. Ciais, N. Viovy, N. de Noblet-Ducoudré, X. Wang, V. Magliulo, M. Wattenbach, L. Vitale, P. Di Tommasi, E. J. Moors, W. Jans, J. Elbers, E. Ceschia, T. Tallec, C. Bernhofer, T. Grünwald, C. Moureaux, T. Manise, A. Ligne, P. Cellier, B. Loubet, E. Larmanou, and D. Ripoche. ORCHIDEE-CROP (v0), a new process-based agro-land surface model: model description and evaluation over Europe. *Geoscientific Model Development*, 9(2):857–873, March 2016.

[215] Florian Zabel, Ruth Delzeit, Julia M. Schneider, Ralf Seppelt, Wolfram Mauser, and Tomáš Václavík. Global impacts of future cropland expansion and intensification on agricultural markets and biodiversity. *Nature Communications*, 10(1):2844, December 2019.

[216] Florian Zabel, Christoph Müller, Joshua Elliott, Sara Minoli, Jonas Jägermeyr, Julia M. Schneider, James A. Franke, Elisabeth Moyer, Marie Dury, Louis Francois, Christian Folberth, Wenfeng Liu, Thomas A.M. Pugh, Stefan Olin, Sam S. Rabin, Wolfram Mauser, Tobias Hank, Alex C. Ruane, and Senthold Asseng. Large potential for crop production adaptation depends on available future varieties. *Global Change Biology*, 27(16):3870–3882, August 2021.

[217] Florian Zabel, Birgitta Putzenlechner, and Wolfram Mauser. Global Agricultural Land Resources – A High Resolution Suitability Evaluation and Its Perspectives until 2100 under Climate Change Conditions. *PLoS ONE*, 9(9):e107522, September 2014.

[218] Yajie Zhang, Yanfen Wang, and Haishan Niu. Spatio-temporal variations in the areas suitable for the cultivation of rice and maize in China under future climate scenarios. *Science of The Total Environment*, 601-602:518–531, December 2017.

[219] Yaojun Zhang, Wenkai Shou, Carmelo Maucieri, and Feng Lin. Rainfall increasing offsets the negative effects of nighttime warming on GHGs and wheat yield in North China Plain. *Scientific Reports*, 11(1):6505, December 2021.

[220] Chuang Zhao, Bing Liu, Shilong Piao, Xuhui Wang, David B. Lobell, Yao Huang, Mengtian Huang, Yitong Yao, Simona Bassu, Philippe Ciais, Jean-Louis Durand, Joshua Elliott, Frank Ewert, Ivan A. Janssens, Tao Li, Erda Lin, Qiang Liu, Pierre Martre, Christoph Müller, Shushi Peng, Josep Peñuelas, Alex C. Ruane, Daniel Wallach, Tao Wang, Donghai Wu, Zhuo Liu, Yan Zhu, Zaichun Zhu, and Senthold Asseng. Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences*, 114(35):9326–9331, August 2017.

APPENDIX

S1 Supplemental Materials for Chapter 2

S1.1 Cultivation Areas

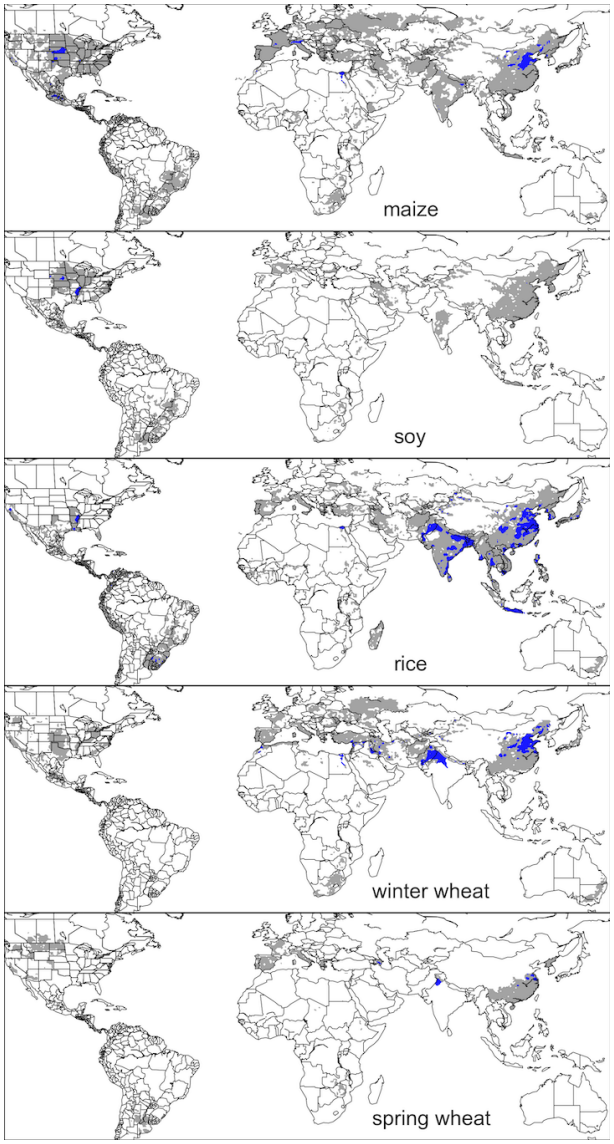


Figure S1: Presently cultivated area for irrigated crops in the real world. The blue contour area indicates grid-cells with more that 20,00 hectares of crop cultivated. The gray contour shows area with more that 10 hectares cultivated. Data from the MIRCA2000 data set for maize, rice, and soy. Winter and spring wheat areas are adapted from MIRCA2000 data and sorted by growing season.

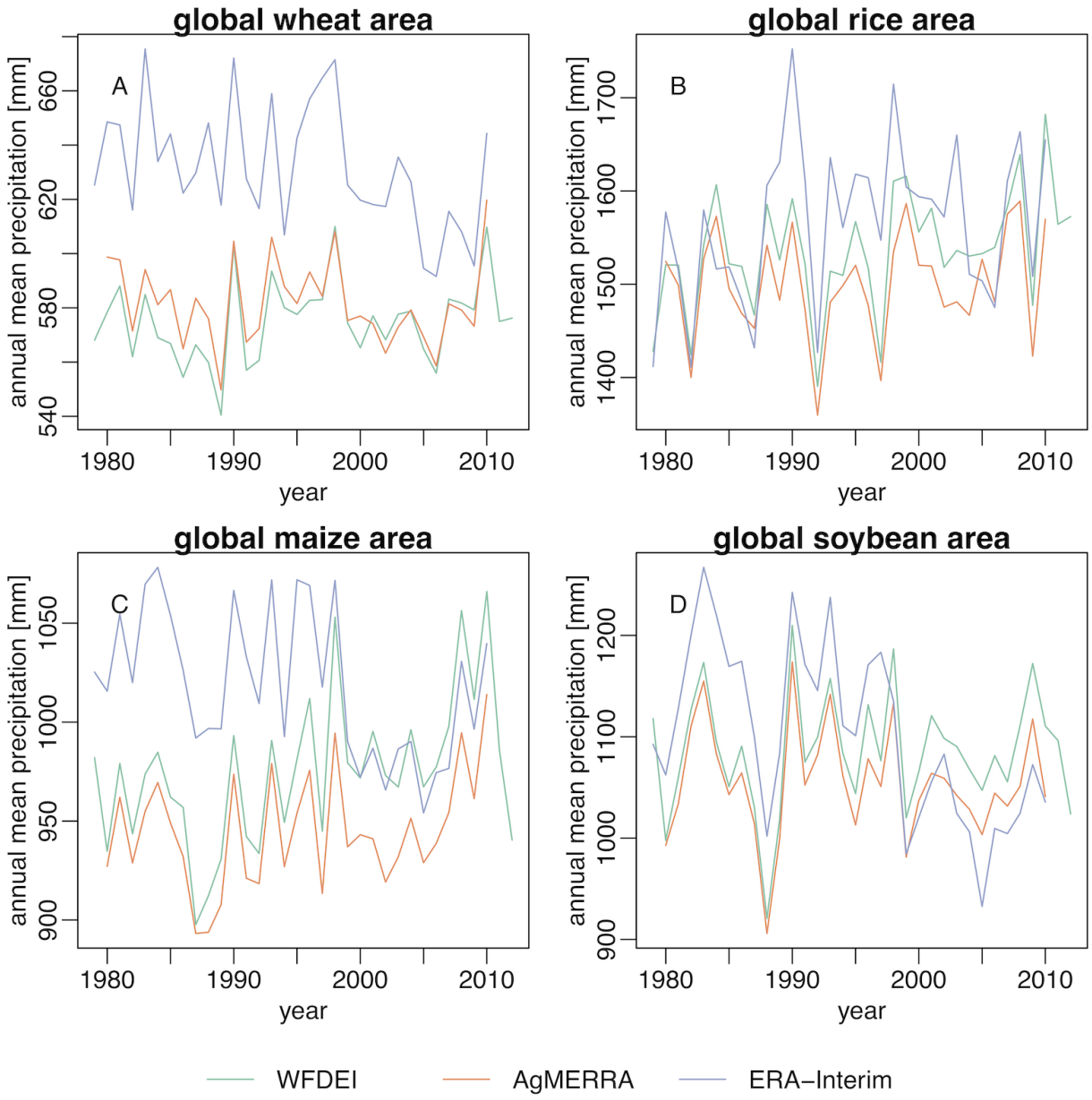


Figure S2: Comparison across the three reanalysis products used in GGCMI Phase 2. Values are aggregated across cultivation area based on the MIRCA2000 dataset.

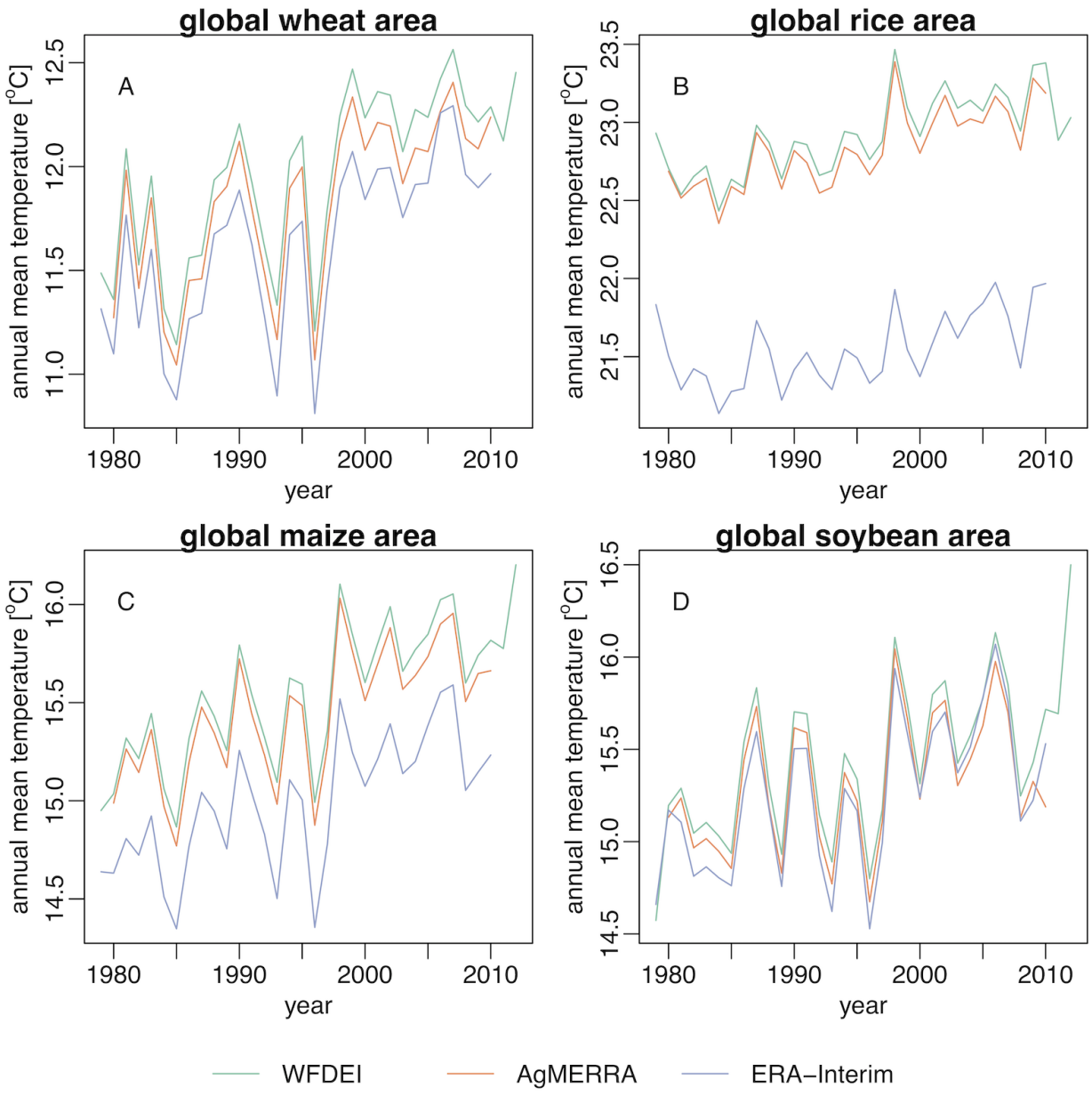


Figure S3: Same as Figure S3 but for temperature.

S1.3 Model Details

Table S1: Key model details. Notes: (NA where not applicable). **a:** D: daily time-step; M: monthly time-step; H: hotextity time-step; WG: use monthly climate data interpolated to daily using a weather-generator. **b:** Ta: average temperature, Tmn: minimum temperature, Tmx: maximum temperature, cld: percentage of cloud cover, sun: fraction of sunshine hours; RH: relative humidity; WS: wind speed; Vap: vapour pressure, Rad: radiation. **c:** Source of soil property inputs (e.g., source of basic soil properties), plus method for manipulation to derive parameters required by the model); AWC: Available Water Capacity 141; HYD: hydraulic soil parameters; THM: thermal parameters; HWSD: Harmonized world soil database 142; STC: soil texture classification based on the USDA soil texture classification (http://ufdc.ufl.edu/IR00003107/00001); ISRIC-WISE 143; ROSETTA 144. **d:** Number of years for spin up (x); OM: organic matter, C: organic carbon; N: organic nitrogen; NH3: ammonia; NO3: nitrate; H2O: soil water; P: phosphorus; CR: crop residues; Tsoil: soil temperature. **e:** calibration of model parameters other than the ones described in the original model description. **f:** PHU+V: prescribed externally computed phenological heat unit requirements and vernalization (winter wheat) per crop and grid cell to meet prescribed harvest date on average (1980-2010); HI: harvest index. **g:** Irrigation rules: depth of soil moisture measured (cm) / lower soil moisture threshold to trigger irrigation (%); / upper soil moisture threshold to stop irrigation (%); / irrigation application efficiency (%); no WS: no water stress. **h:** Irrigation rules: EPIC-based models: water stress in crop to trigger automatic irrigation (%); / irrigation efficiency - runoff from irrigation water (%); / maximum of annual irrigation volume (mm); / maximum of single irrigation volume allowed (mm); / minimum of single irrigation volume allowed (mm). **i:** Remove residue or not (Yes/No). **j:** ET0: LSM: land surface model, complex computation of energy and water vapor fluxes

Model	Temporal scale (a)	Climate input variables (b)	Soil input data (c)	Spin-up (d)	specific calibration other than growing season (e)	Crop cultivars (f)	Irrigation rules (g,h)	Crop residue removal (i)	model category	primary production	ET0 (j)
APSIM-UGOE	D	Tmn, Tmx, P, Rad	ISRIC-WISE	Soil OM, C, NH3, NO3, H2O (1)	co2 effects included for maize and soybean	N/A	120/100/no/no water stress	Yes	field scale	Radiation use efficiency	Priestley-Taylor
CARAIB	D, but 2-hourly for photosynthesis	Tmn, Tmx, P, Rad, RH,WS	HWSD	Soil OM, C, H2O, Tsoil (10)	acclimation of photosynthesis to CO2 and acclimation of autotrophic respiration to temperature included	PHU, HI	80-125*/100/100/100 (g) (*) species-dependent root depth	Yes	global ecosystem	photosynthesis	Penman
EPIC-IIASA	D	Tmn, Tmx, P, Rad, RH, WS	ISRIC-WISE; ROSETTA; AWC; HYD 137	Soil OM, C, NH3, NO3, H2O, P, CR (20)	none	PHU+V	90/100/2000/50/0/0 (h)	No	field scale	Radiation use efficiency	Hargreaves
EPIC-TAMU	D	Tmn, Tmx, P, Rad, RH, WS	ISRIC-WISE	Soil OM, C, NH3, NO3, H2O, P, CR (10)	N/A	PHU+V	99/100/9999/100/25 (h)	No	field, region, global	Radiation use efficiency	Penman-Monteith
GEPIC	D	Tmn, Tmx, P, Rad, RH, WS	HWSD	Soil OM, C, NH3, NO3, H2O, P, CR (20)	none	PHU+V, HI (mai)	99/100/2000/1000/0 (h)	Yes, Crop-specific	field scale	Radiation use efficiency	Hargreaves
JULES	3-hourly	T,rain,snow,ShortRad,LongRad, WS,specific humidity, pressure	HWSD	H20, Tsoil (110)	none	PHU	35/-10kPa/-10kPa/100 (g)	Yes	global ecosystem	photosynthesis	LSM
LPJ-GUESS	D	Ta, P, Rad	ISRIC-WISE	C, N-organic, N-inorganic, CR, Tsoil (500 grassland)	none	PHU+V	200/90/100/100 (g)	Yes, 10% left on field from previous cycle	global ecosystem	photosynthesis	Priestley-Taylor
LPJmL	D	Ta, P, cld (or ShortRad, LongRad)	HWSD, STC HYD138, THM 139	C, N, NH3, NO3, H2O, CR, Tsoil (5000 natural vegetation, 390 land use)	none	PHU+V	300/90/100/varies (g)	Yes, 10% left on field from previous cycle	global ecosystem	photosynthesis	Priestley-Taylor
ORCHIDEE-crop	Half-hourly	Tmn, Tmax, P, Rad, RH, WS	NA	H2O (1)	none	PHU+V	200/90/100/varies (g)	Yes, does not affect yield	global ecosystem	photosynthesis	Penman-Monteith
pDSSAT	D	Tmn, Tmx, P, ShortRad	Global Soil Dataset for Use in Earth System Models	Soil OM, C, NH3, NO3, H2O (1)	none	PHU+V (for anthesis and maturity)	mai and soy: 40/99/100/100 ric, swh, wwh: 40/80/100/100	Yes, does not affect yield	field scale	Radiation use efficiency	Priestley-Taylor
PEPIC	D	Tmn, Tmx,P, Rad, RH, WS	ISRIC-WISE	Soil OM, C, NH3, NO3, H2O, P, CR (20)	none	PHU+V	90/100/1000/500/1 (h)	Yes	field scale	Radiation use efficiency	Penman-Monteith
PROMET	H	T,P,ShortRad,cld,LongRad,WS, RH,P	Derived from HWSD	H2O(5), Tsoil(5)	none	PHU+V	no WS	No	global ecosystem	photosynthesis	LSM

Calibration procedures for growing season re-calibration in the A1 scenarios:

- APSIM-UGOE: Default cultivars with forced harvest at given maturity days.
- CARAIB: Harvest forced on the same day as in the A0 simulation.
- EPIC-IIASA: potential Heat Units (PHU) were calculated for each crop and grid cell based on grid-specific input sowing and harvesting dates using the background weather dataset
- EPIC-TAMU: An algorithm was written to calculate the heat units required to reach maturity for a particular crop and location. These heat units were calculated for each year in the weather record, and PHU was set as the average heat units from this time series.
- GEPIC: Crop- and pixel-specific potential heat units (PHU) were estimated ex ante based on input sowing dates, harvest dates, and 31 year monthly means of minimum and maximum temperatures using a program provided by the EPIC development team. The resulting long-term average PHU were subsequently used as a model input parameter.
- JULES: Thermal times from emergence to flowering and flowering to harvest, were created by iteratively working out which thermal times would produce the right harvest dates for each crop in each gridbox (using the sowing and maturity dates provided by GGCMI and calculating harvest dates using table 11 in the Phase 1 protocol) using a 3hrly WFDEI climatology for the years 1991-2000 inclusive. The thermal time between emergence and flowering was assumed to be a crop-specific constant fraction of the thermal time between emergence and harvest. These thermal times were then prescribed as crop- and grid-cell specific values.

- LPJ-GUESS: In a preparatory simulation run with unlimited growing season length, the accumulated phenological heat units (PHU) at the given maturity date were recorded per crop, year and grid cell. These heat units were averaged over all years and then prescribed as crop- and grid-cell specific values.

- LPJmL: In a preparatory simulation run with unlimited growing season length, the accumulated phenological heat units (PHU) at the given maturity date were recorded per crop, year and grid cell. These heat units were averaged over all years and then prescribed as crop- and grid-cell specific values.

- ORCHIDEE-crop: Crop-specific thermal times from emergence to flowering and flowering to harvest were created from default datasets of cultivars, by a testing simulation during 2000s, we chose the cultivar best matching the calendar provided by the protocol.

- pDSSAT: Ran a calibration simulation in the baseline period, proportionally adjusting the phenological GDD parameters (p1 and p5 for grains) to produce thr target average growing season length over the baseline period. The resulting calibrated parameters were used for future simulations.

- PEPIC: Used averaged month temperature across the study period, prescribed growing season, and crop-specific base temperature to estimate PHU for each grid and crop.

- PROMET: In a preparatory sensitivity analysis, simulation runs with unlimited growing season length were carried out for different cultivars. A 'cultivar factor' was set to a crop, year and grid cell specific value that reproduces the statistical growing season. This 'cultivar factor' was averaged over all 30 years and then prescribed as crop- and grid-cell specific values.

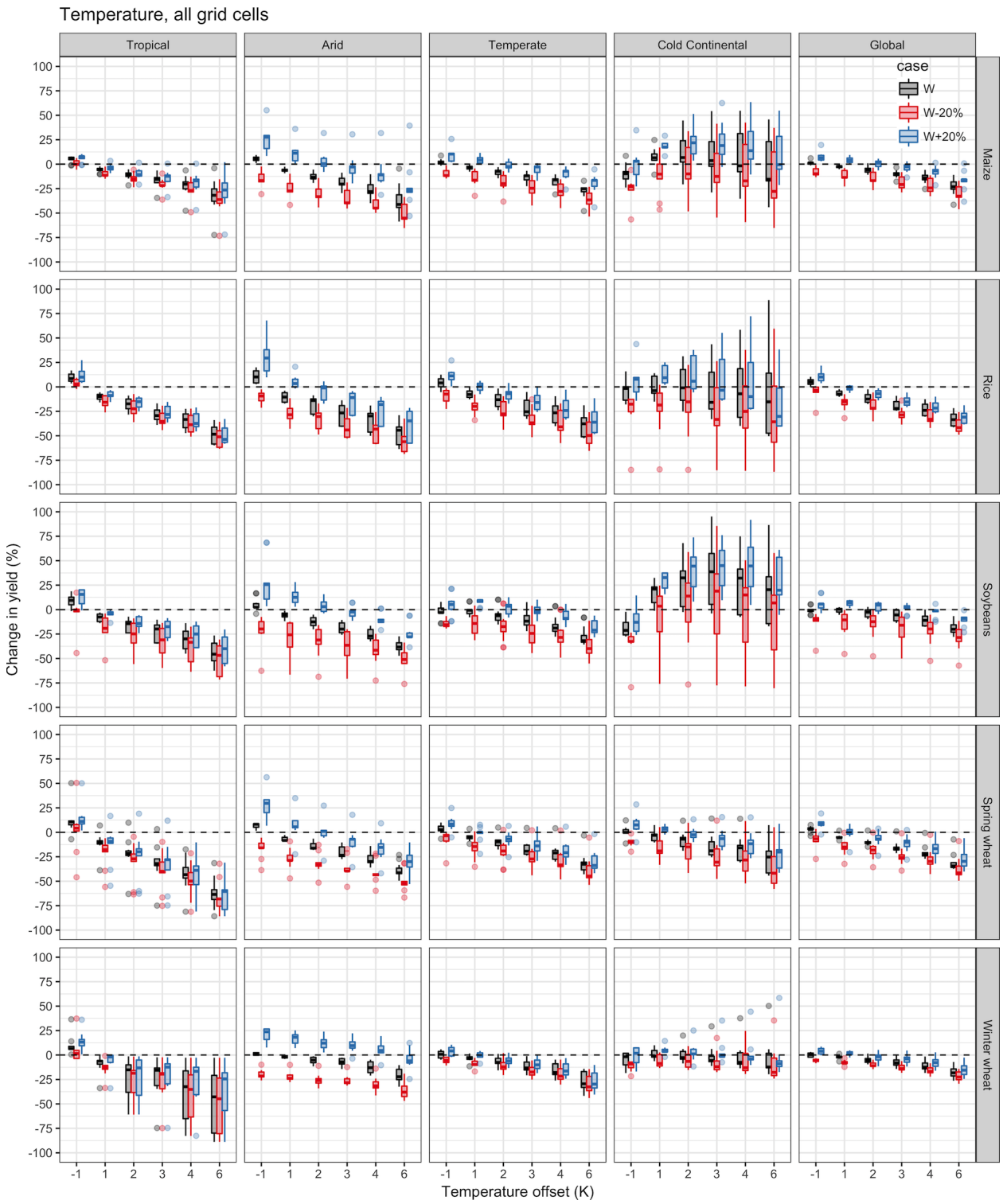


Figure S4: Same as Figure 5a for all crops.

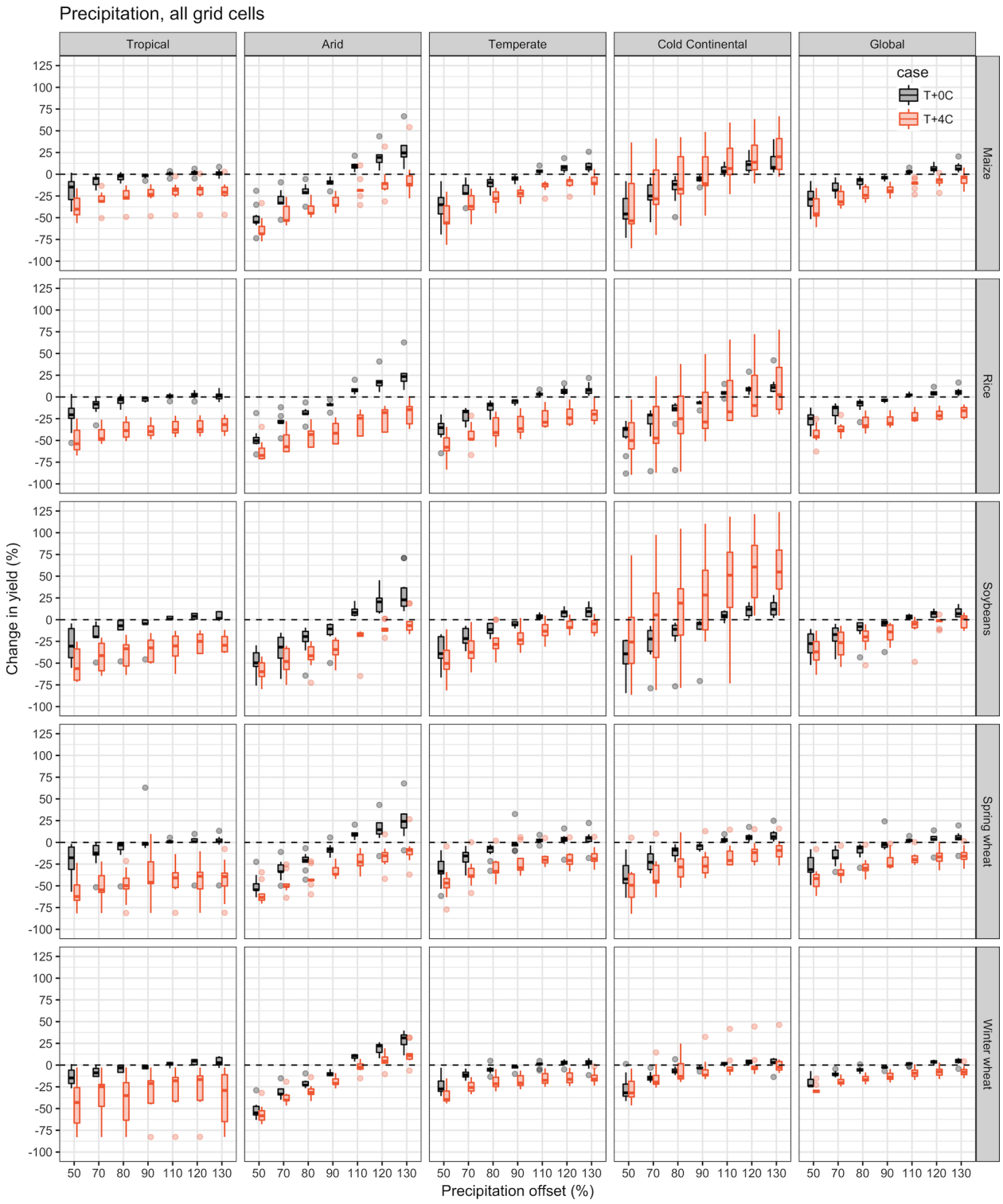


Figure S5: Same as Figure 5b for all crops.

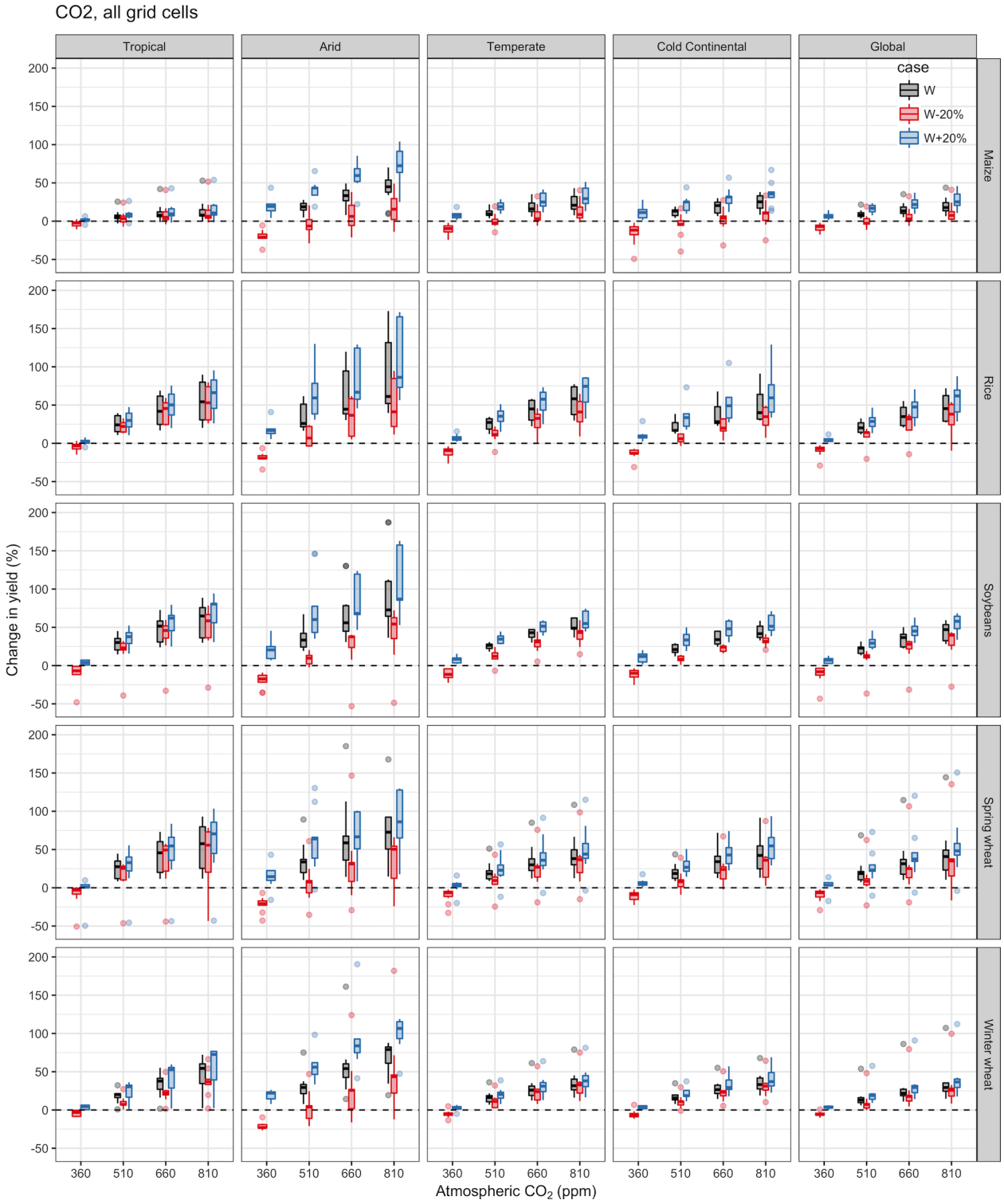


Figure S6: Same as Figure 2.6a for all crops.



Figure S7: Same as Figure 2.6b for all crops.

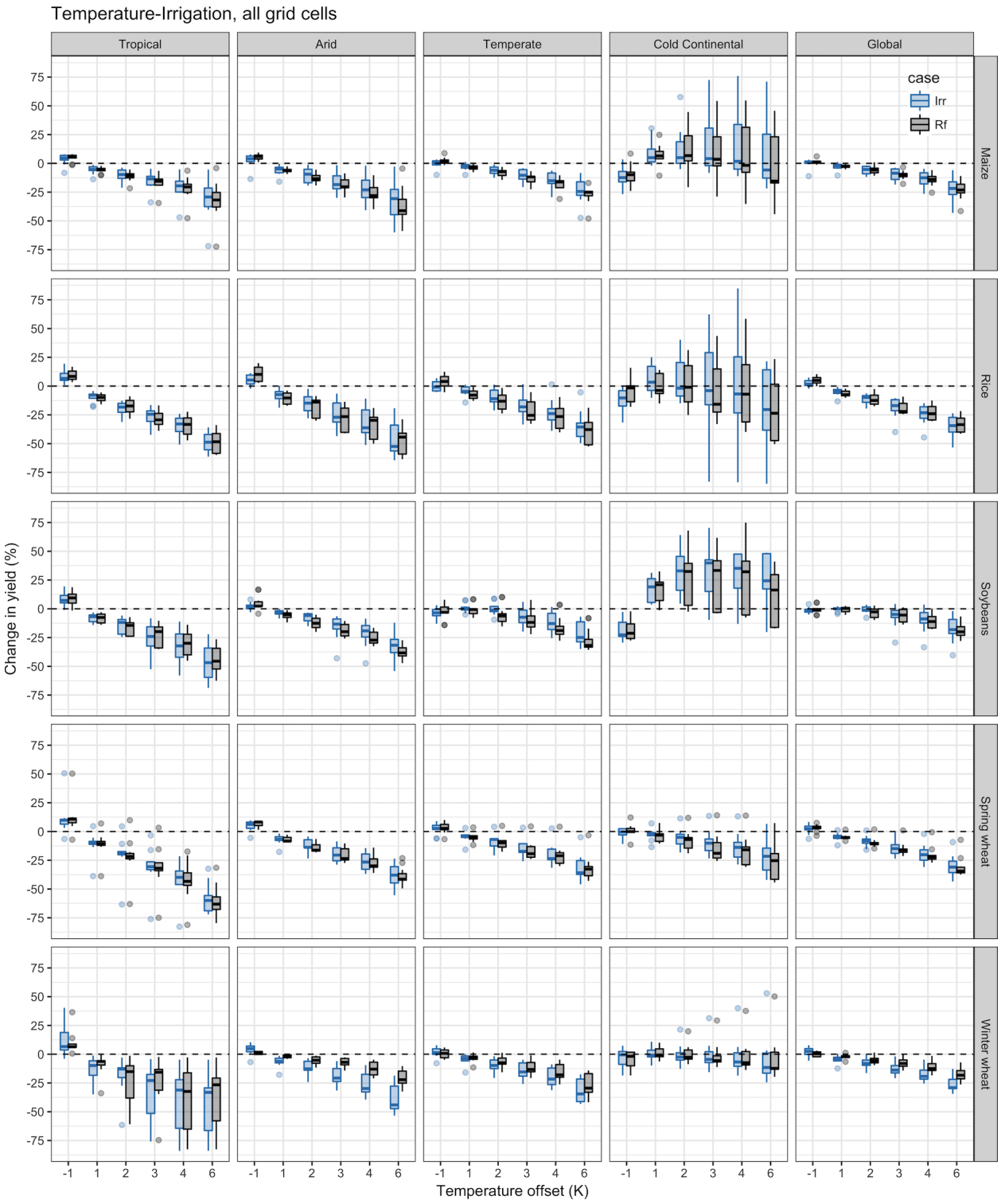


Figure S8: Same as Figure 2.5a for all crops. Irrigated crops compared to rainfed. Note that yield change for irrigated crops is from the irrigated baseline, which is typically higher than rainfed.

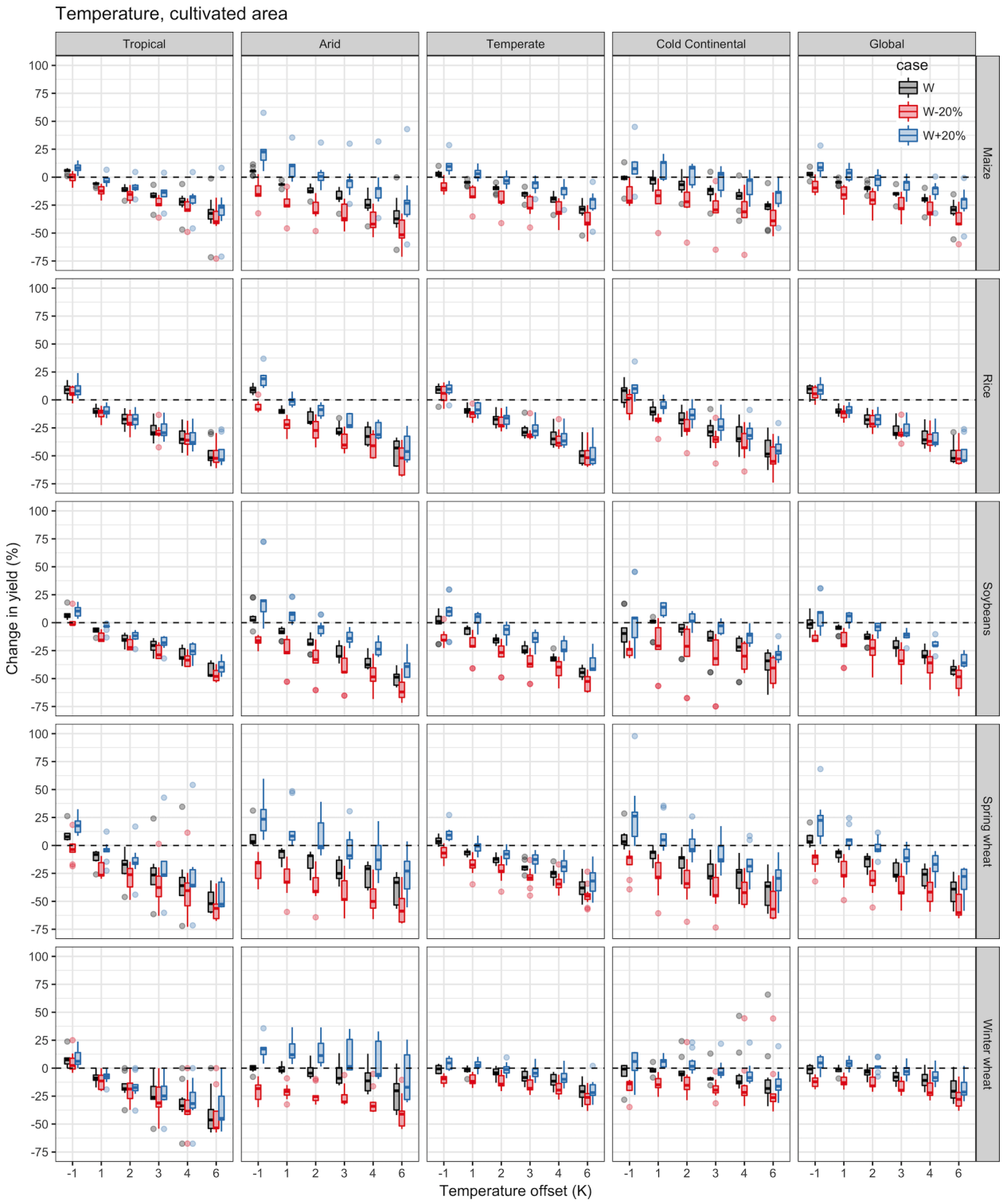


Figure S9: Same as Figure 2.5a for all crops. Only over cultivated area.

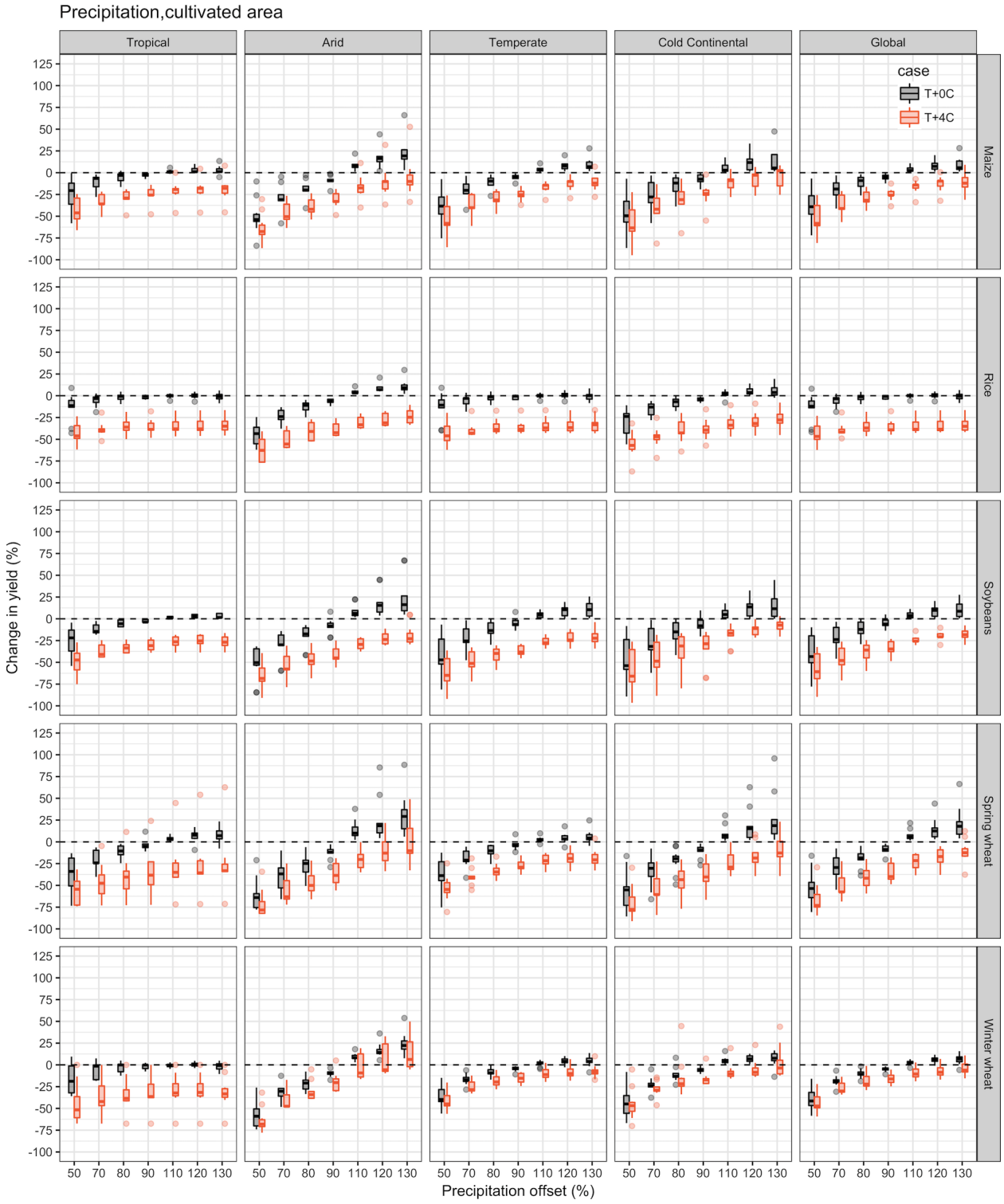


Figure S10: Same as Figure 2.5b for all crops. Only over cultivated area.

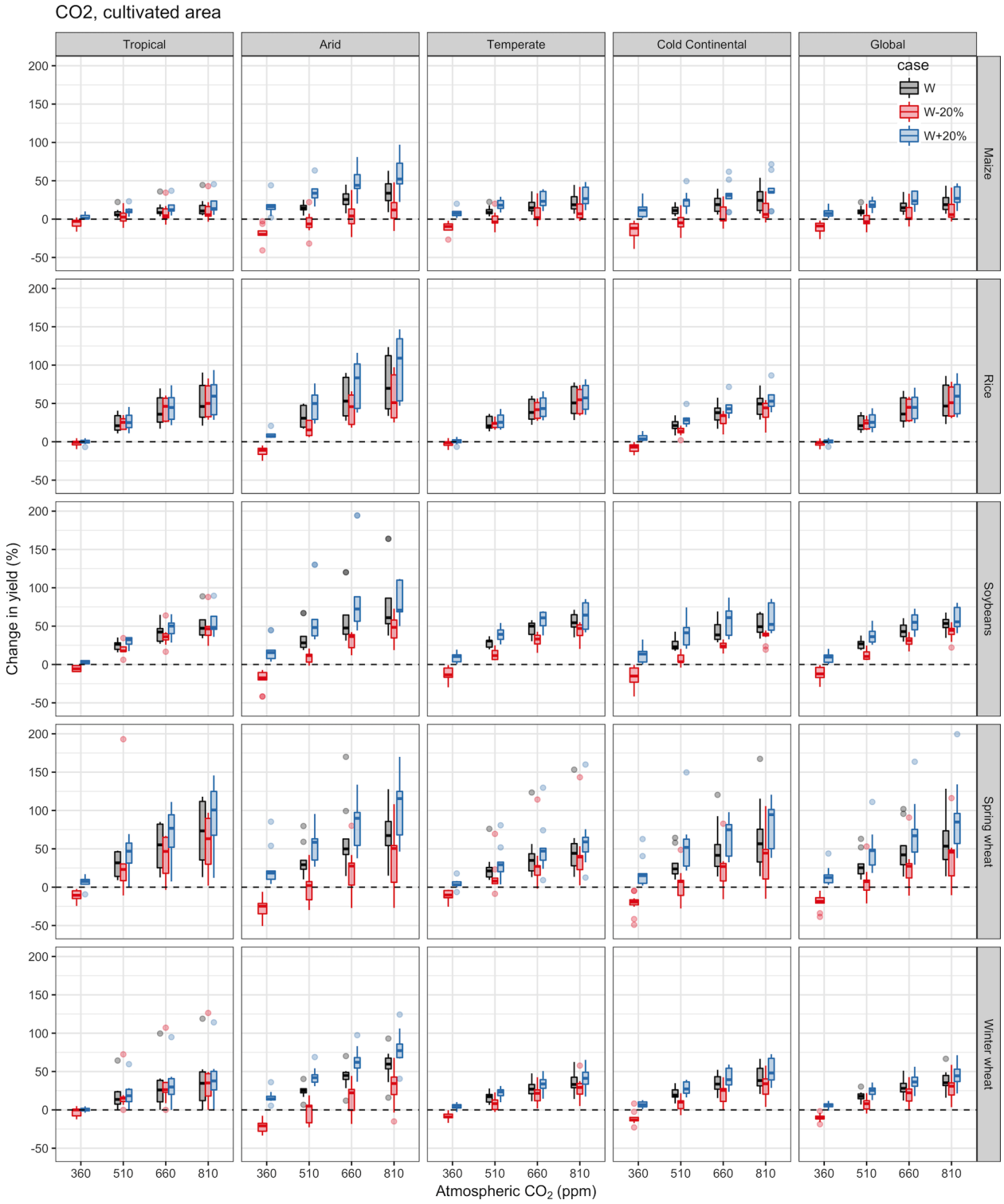


Figure S11: Same as Figure 2.6a for all crops. Only over cultivated area.

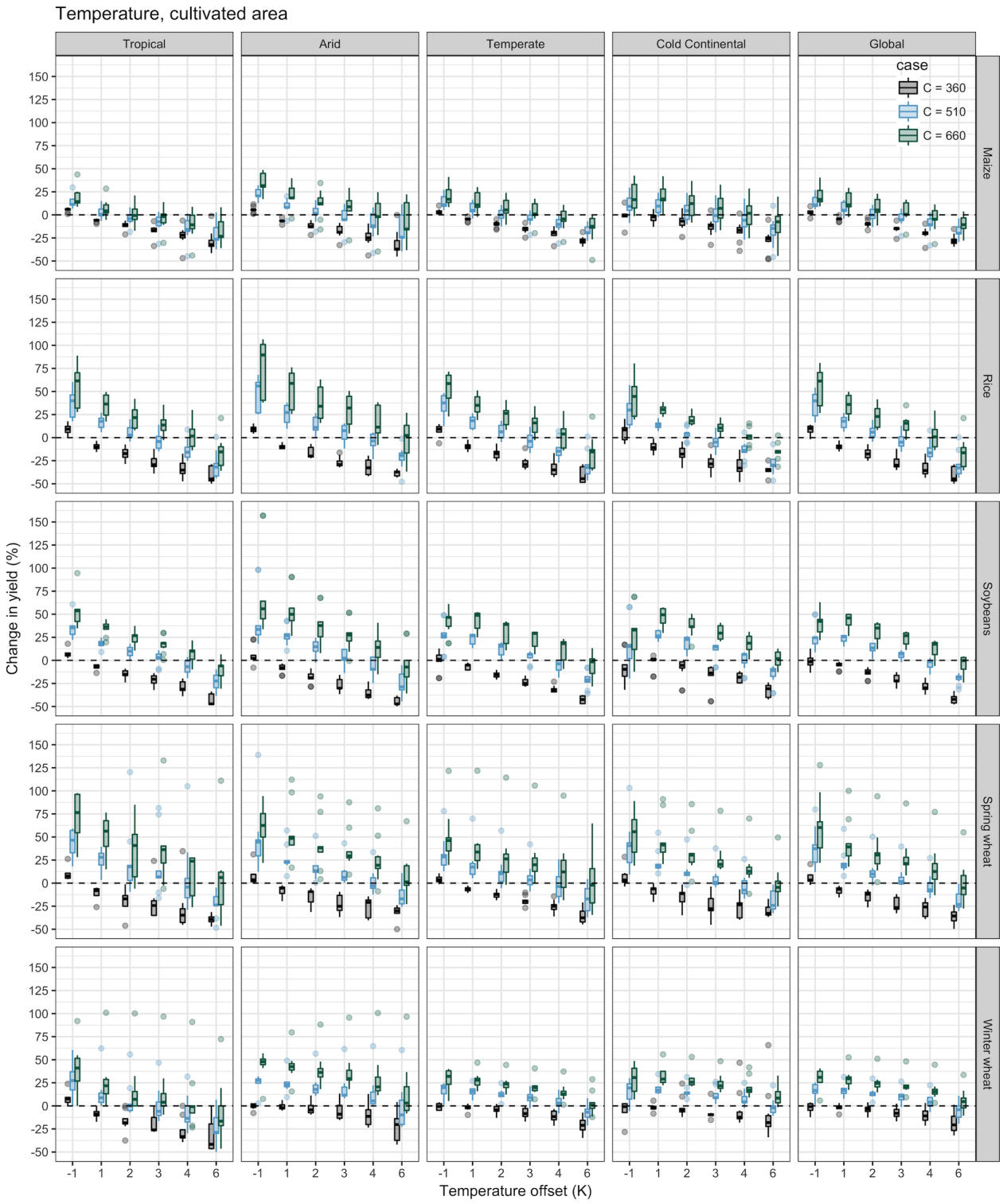


Figure S12: Same as Figure 2.6b for all crops. Only over cultivated area.

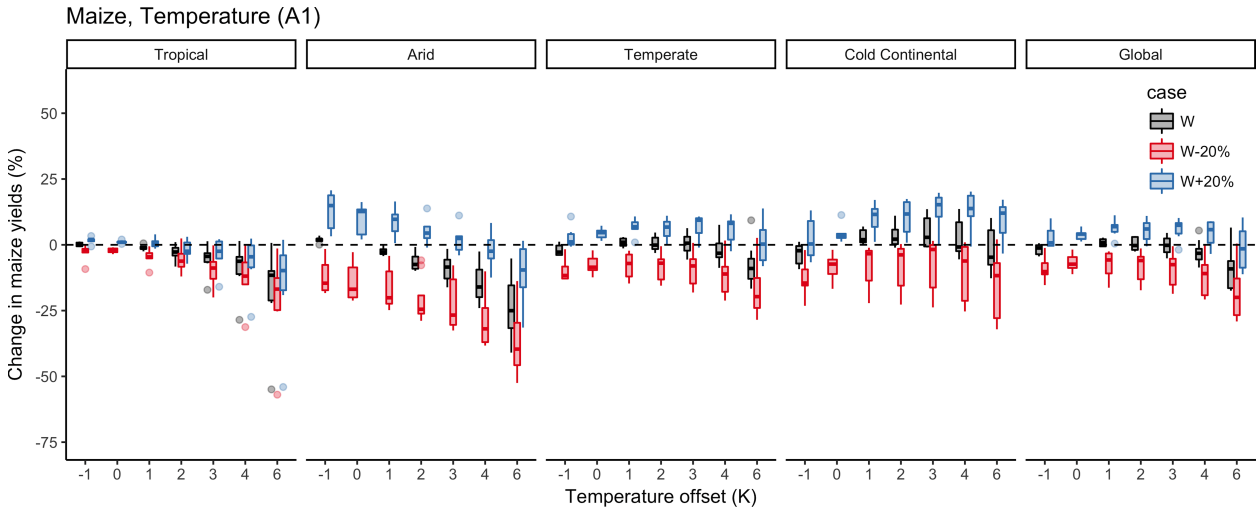


Figure S13: Same as Figure 2.5a for except for A1 simulations where the growing season in held constant under warming.

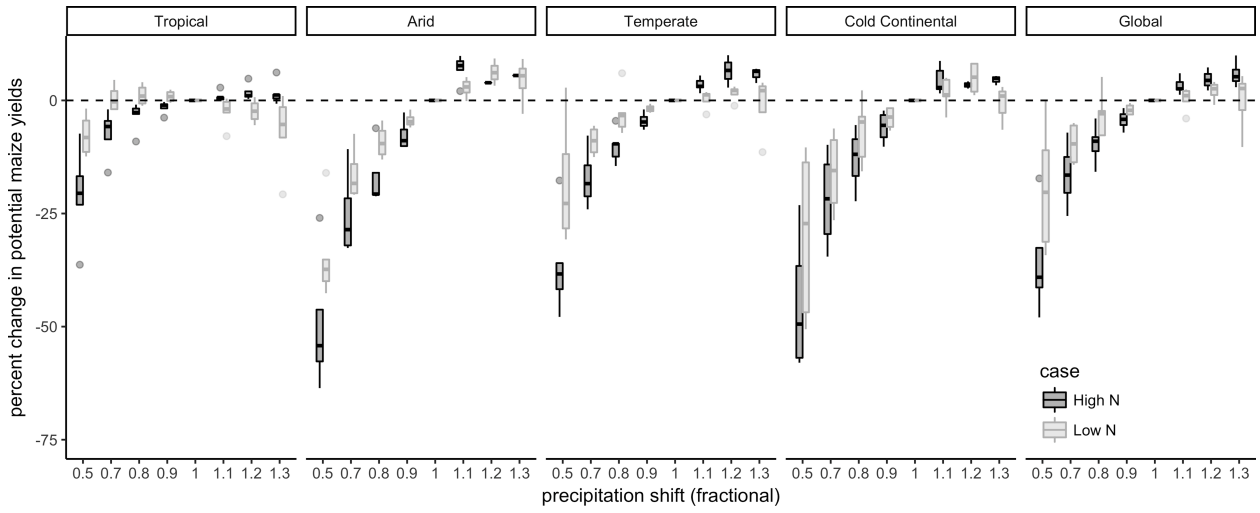


Figure S14: Same convention as Figure 2.5b except for maize across the precipitation and nitrogen dimensions.

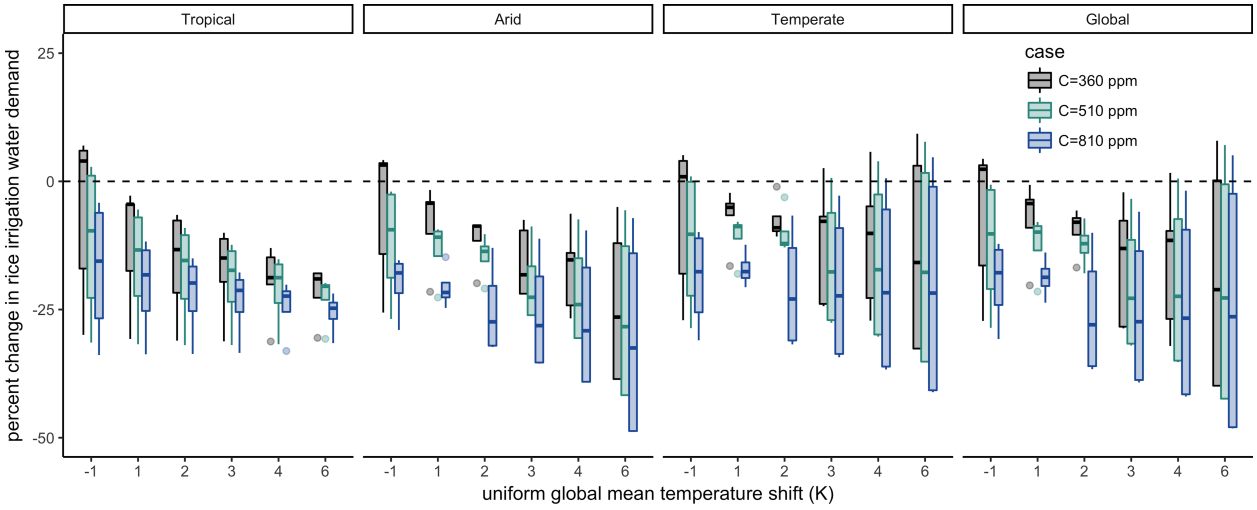


Figure S15: Same convention as Figure 2.6b except for irrigation water demand instead of yield.

S2 Supplemental Materials for Chapter 3

S2.1 Sampling

Simulation sampling across the defined variable space is not uniform in the GGCMI Phase 2 experiment, with only some models providing all cases in the protocol. Figure S16 compares the sampling density of the models used in the emulator analysis.

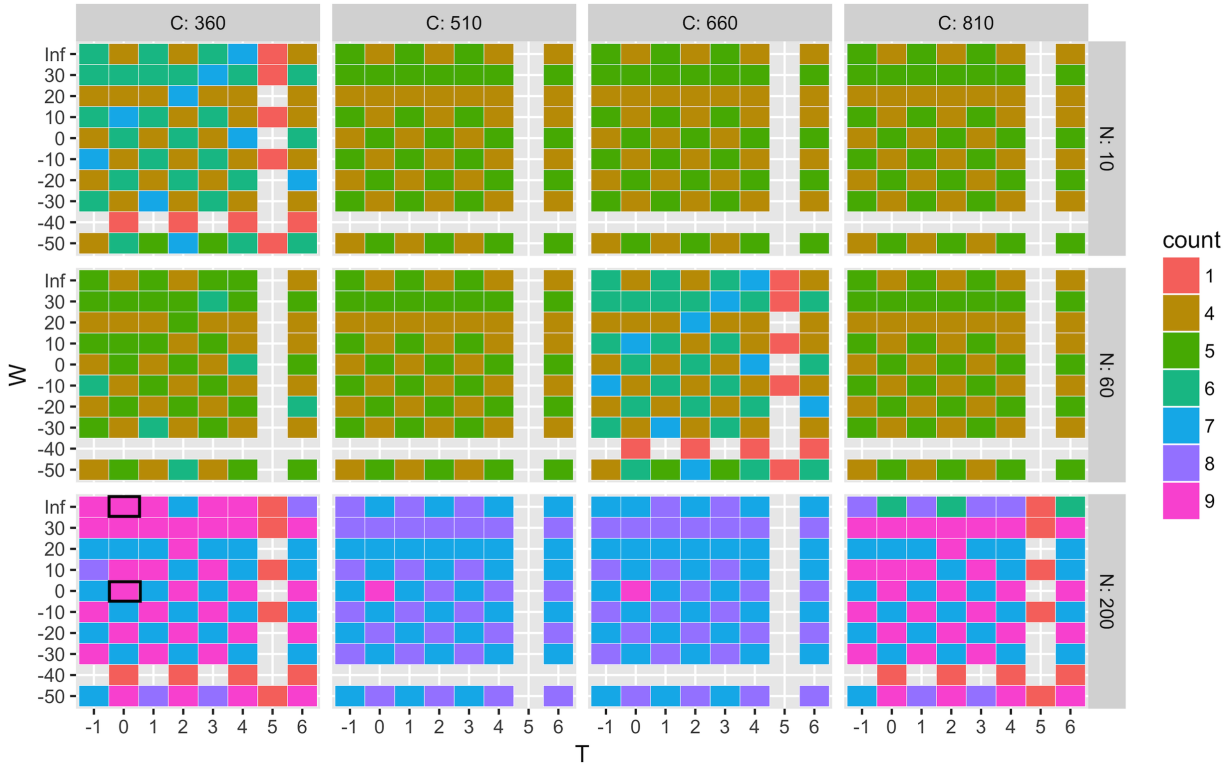


Figure S16: Heatmap illustrating the number of models providing simulations for each of the scenarios in CTWN variable space. Black boxes mark the "baseline" cases for rainfed and irrigated simulations.The maximum number is 9, the number of models included in the emulator analysis. (That is, we exclude here the three GGCMI Phase II models not included in the emulator analysis.) For cases with N levels lower than 200 kg/ha, the maximum number of models is 6 since three models (CARAIB, JULES, and PROMET) do not represent varying N levels. One model (GEPIC) provided additional simulations at T+5 not specified by the protocol; these are not used in emulation.Normalized error calculations are run only over scenarios in which 9 models contribute simulations (pink boxes).

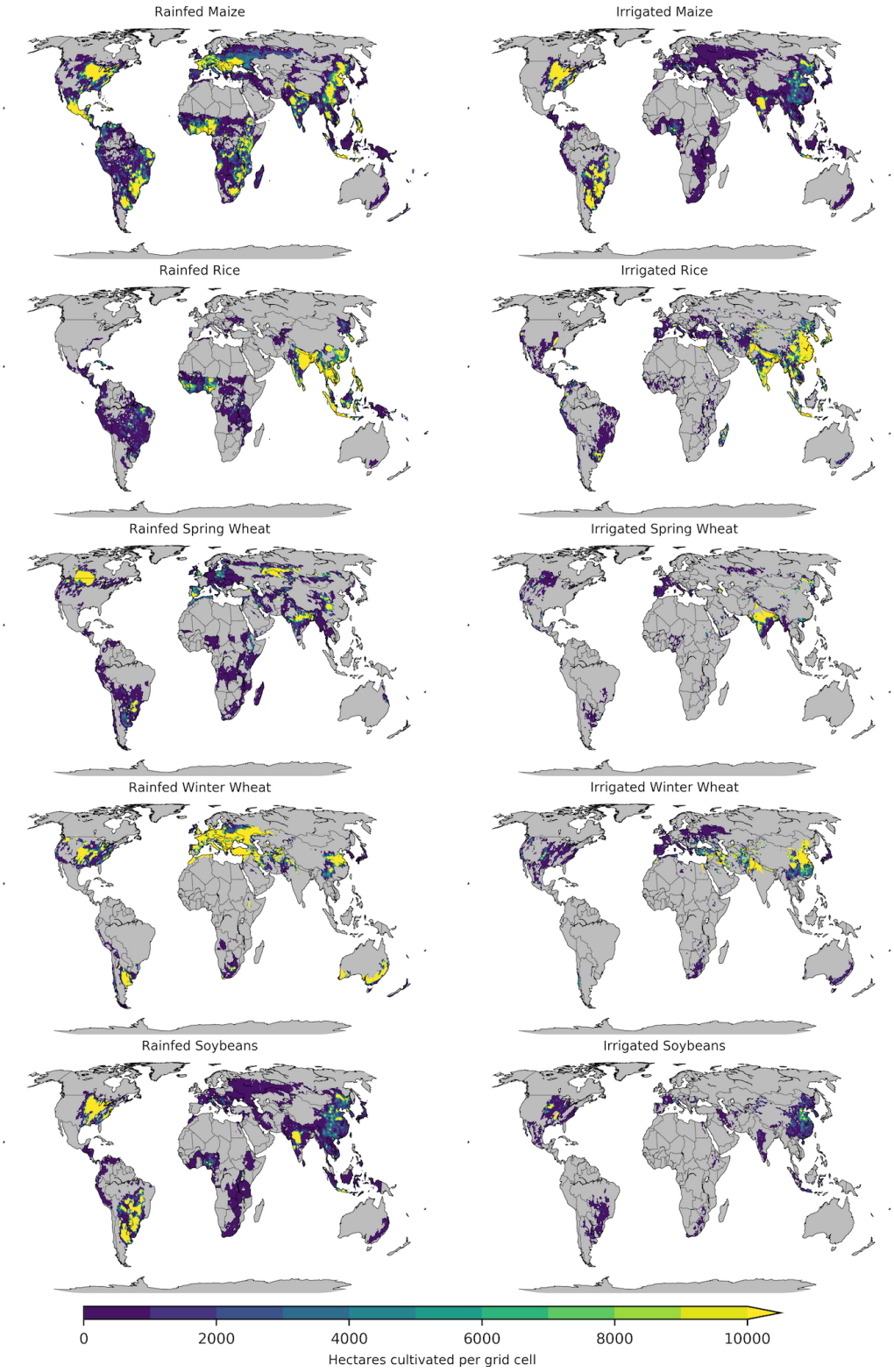


Figure S17: Presently cultivated area in the real world for rainfed (left) and irrigated (right) crops, from the MIRCA2000 dataset. Blue areas show grid-cells with more that 20,00 hectares of crop cultivated, and gray areas those with more that 10 hectares cultivated. Data are taken directly from the MIRCA2000 dataset for maize, rice, and soy. Winter and spring wheat areas are adapted from MIRCA2000 and sorted by growing season.

S2.2 Variability Changes in Future Climate Projections

Because the GGCMI Phase 2 simulation dataset does not sample across changes in climate variability, large impacts to yields driven by future changing variability would decrease the practical utility of the emulator for impacts assessments. We therefore assess the scale of potential future changes in temperature variability, in RCP8.5 simulations from the five climate models used in ISIMP. In chapter 2 we use one of these climate simulations, that from HadGEM2-ES, to assess the ability of GGCMI emulators to reproduce yield changes simulated under more realistic climate projections. We choose the HadGEM2-ES model because it shows the largest variability changes, and therefore provides a stricter test of the utility of a GGCMI emulator. Table S2 summarizes daily T_{max} variability changes for each crop and model weighted by production. Figures S19 and S20 below show changes in variability in minimum and maximum temperatures in the HadGEM2 simulation for each crop growing season and area, and Figure S5 shows changes in daily T_{max} variability for maize across the 4 additional ISIMIP climate simulations. (Compare to Figure S19 upper left panel.) Most crop models included in GGCMI Phase 2 take daily minimum and maximum temperature as inputs, though PROMET and JULES take sub-daily temperature inputs.

Table S2: Global production-weighted fractional change in growing season daily maximum temperature variability under RCP8.5 for the five climate models included in the ISIMIP project (a subset of the CMIP-5 archive). Value for each crop and model is mean within-growing season temperature standard deviation across 30 growing seasons of 2070-2099 relative to that for 1981-2010, with grid-cell values weighted by LPJmL model simulated yields and current cultivation area (MIRCA).Values in parenthesis are the change in variability by the same metric for daily minimum temperature within the growing season.The HadGEM2-ES model is highlighted in bold because this model is used for our emulator evaluation in Chapter 3. The HadGEM2-ES model is chosen because it shows the highest changes in variability.

Model	**Maize %**	**Soybean %**	**Rice %**	**S. Wheat %**	**W. Wheat %**
HadGEM2-ES	**9.7 (2.1)**	**10.4 (-0.6)**	**10.1 (-3.3)**	**6.4 (4.7)**	**3.6 (1.7)**
GFDL-ESM2M	3.6 (0.9)	3.4 (0.6)	2.7 (-0.3)	1.2 (-0.3)	2.0 (1.0)
NorESM1-M	6.7 (-1.1)	6.5 (-4.0)	5.9 (-3.5)	4.5 (3.3)	2.4 (0.8)
IPSL-CM5A-LR	3.3 (3.3)	3.4 (0.6)	3.4 (1.4)	1.4 (2.3)	1.2 (1.3)
MIROC-ESM	4.0 (1.2)	3.1 (0.1)	0.4 (-5.1)	2.2 (-0.3)	2.6 (2.3)

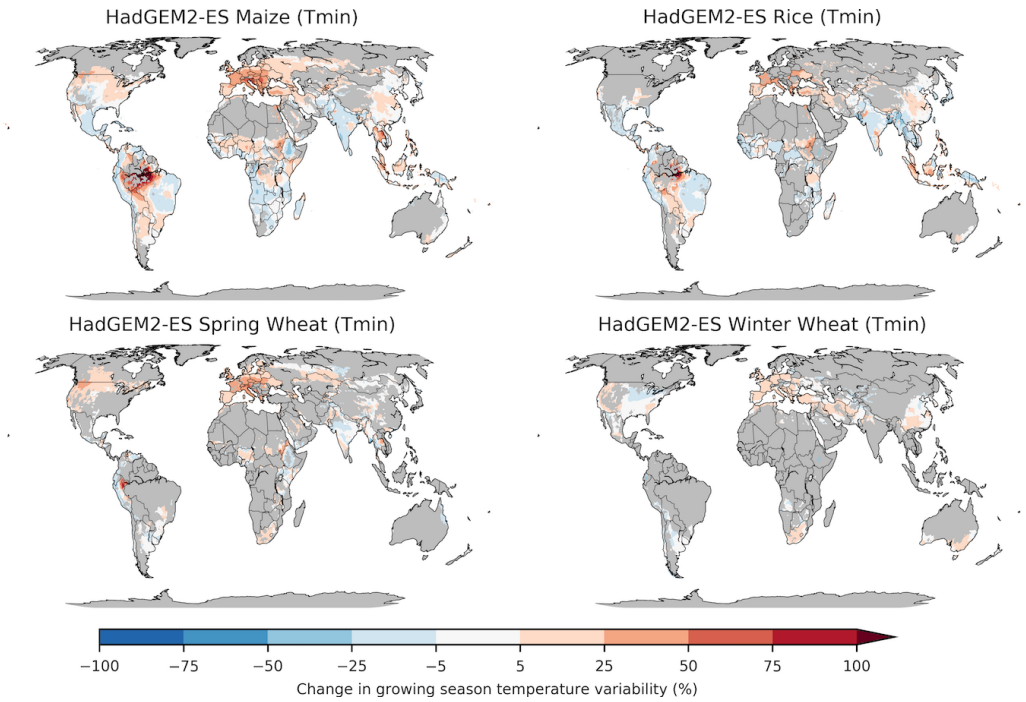


Figure S18: Change in within-growing-season daily minimum temperature variability by the end of the century under RCP8.5 for the HadGEM2 model, computed for selected crops. Heatmap shows the percentage change for 2070-2099 relative to 1981-2010 over currently cultivated areas, with red indicating increases and blue decreases. To determine the change we compute the mean standard deviation of daily T_{min} in each historical growing season and take the mean across all 30 years; this metric therefore includes changes both in seasonality and in short-term variations but excludes inter-annual variability and longer-term trends. For winter wheat, growing-season variability reductions reflect the dampening of the seasonal cycle (stronger warming in winter). Strong percentage increases in the tropics reflect very low variability in the baseline. Production-weighted mean changes across crops range from -3% for rice to +5% for spring wheat (Table S1). Note that changes may differ if calculated using an ensemble of simulations rather than a single projection as is done here.

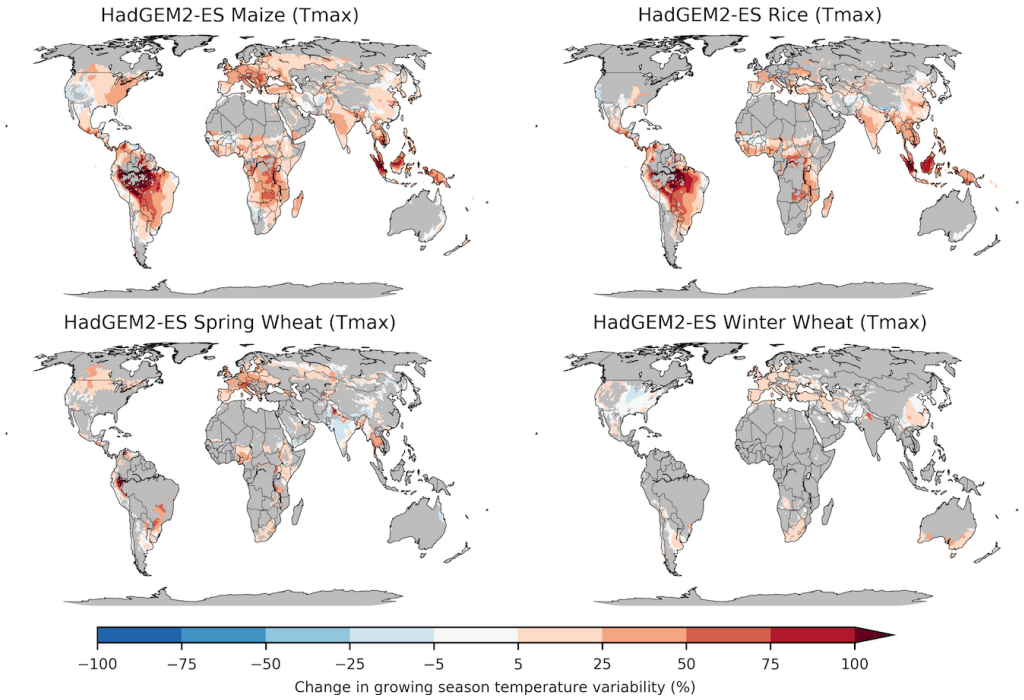


Figure S19: As in Figure S18 except now for daily maximum temperature. Changes in daily maximum temperature variability are generally more positive than those for daily minimum temperature.

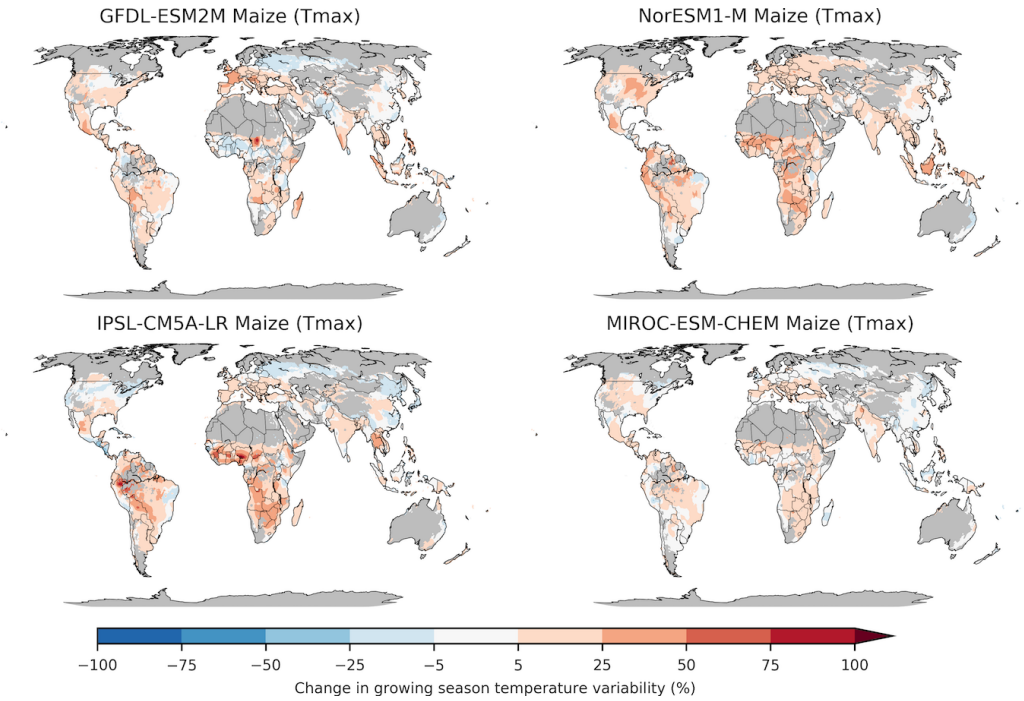


Figure S20: As in Figure S18, change in daily maximum temperature variability, except now for maize only, for the remaining 4 ISIMIP climate simulations. Values are lower on average than for HadGEM2-ES but patterns can differ.

S2.3 Yield Response for A1 (Growing Season Adaptation) simulations

This section shows illustrations of emulator ability to capture yield changes in A1 simulations; compare to Figures 3.5 and 3.6 showing A0 simulations. Responses to CWN factors are similar in both but responses to T are substantially weaker in A1 simulations, in which growing season length does not contract in warmer future conditions.

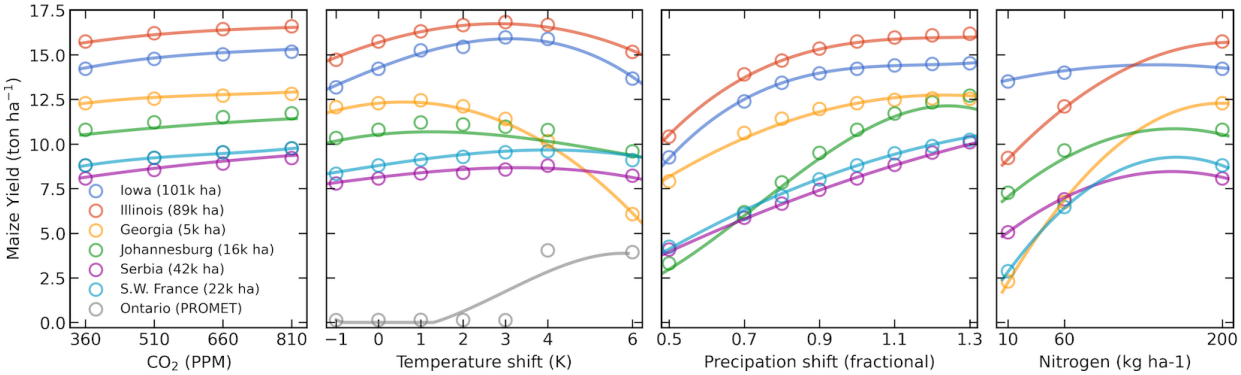


Figure S21: Illustration of spatial variations in yield response, which are successfully captured by the emulator for the A1 simulations. Panels show simulations (points) and emulations (lines) of rainfed maize in the pDSSAT model in six example locations selected to represent high-cultivation areas around the globe. Legend includes hectares cultivated in each selected grid cell. Each panel shows variation along a single variable, with others held at baseline values.

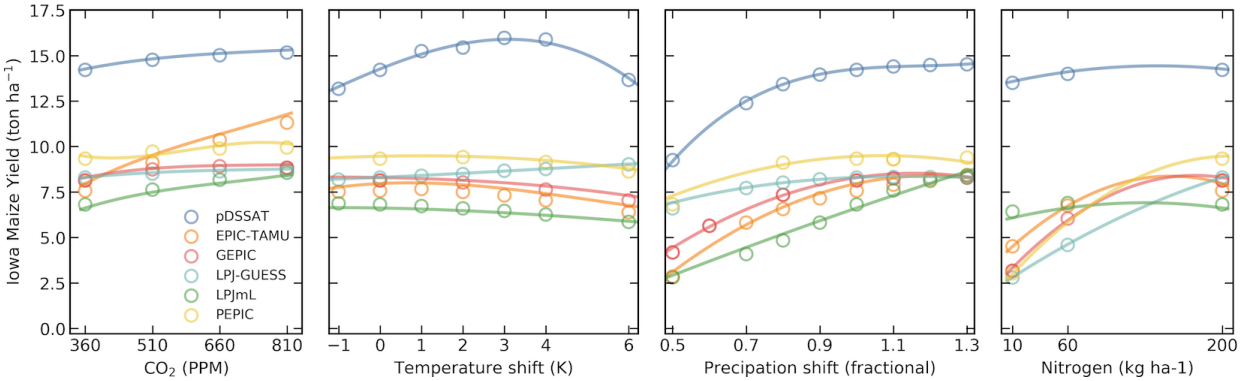


Figure S22: Illustration of variations in yield response across models for A1 simulations, again successfully captured by the emulator. Panels show simulations and emulations from six representative GGCMI models for rainfed maize in the same Iowa grid cell shown above, with the same plot conventions. Three models (PROMET, JULES, and CARAIB) that do not simulate the nitrogen dimension are omitted for clarity.

S2.4 Normalized Error for Other Cases

In Chapter 3, we show normalized error for the A0 emulators over all rainfed crops, models, and T and W values for baseline CO_2 and nitrogen levels (360 ppm and 200 kg ha-1). Here we show normalized error in some alternate cases for comparison: *Figure S6:* A0 emulators of rainfed crops at higher CO_2, *Figure S7:* A1 emulators of rainfed crops at baseline values, *Figure S8):* A0 emulators of irrigated crops at baseline values. Results are generally similar, with a few exceptions. Normalized errors at higher CO_2 are generally lower because model disagreement is larger, lowering the denominator. Some model emulators for irrigation water demand are under-performing: LPJ-GUESS and CARAIB for some crops. A1 errors are larger than A0 errors for several crops and models: LPJmL rice, pDSSAT spring wheat, and PROMET winter wheat.

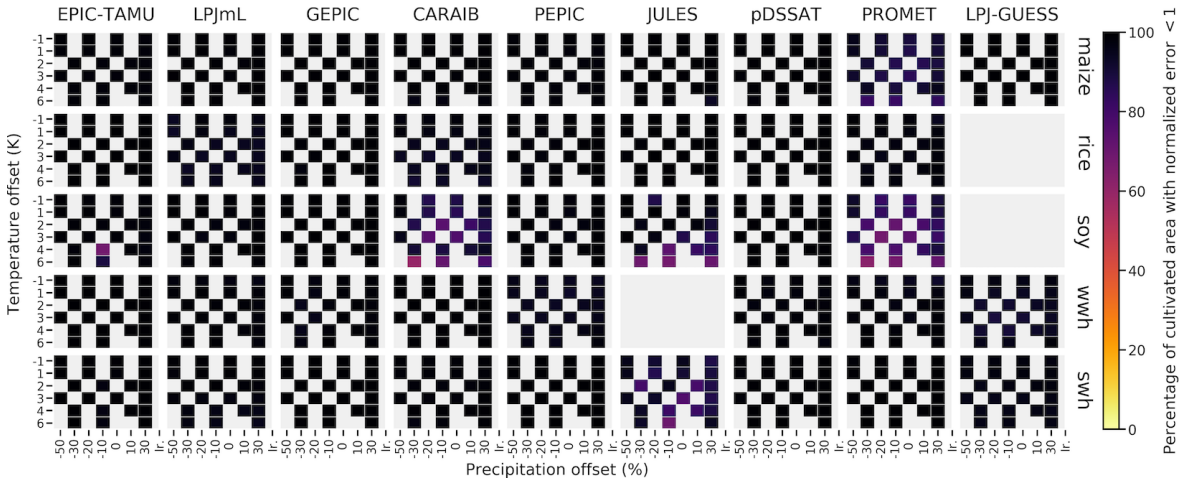


Figure S23: Fraction of currently cultivated hectares with normalized emulation error less than 1 for the CO_2=810 ppm and 200 kg N ha^{-1} yr^{-1} case for the temperature and precipitation perturbations scenarios provided by all 9 models included in the emulator analysis. Figure convention as in Figure 3.7. The yield response is generally easy to emulate over currently cultivated areas (black regions).

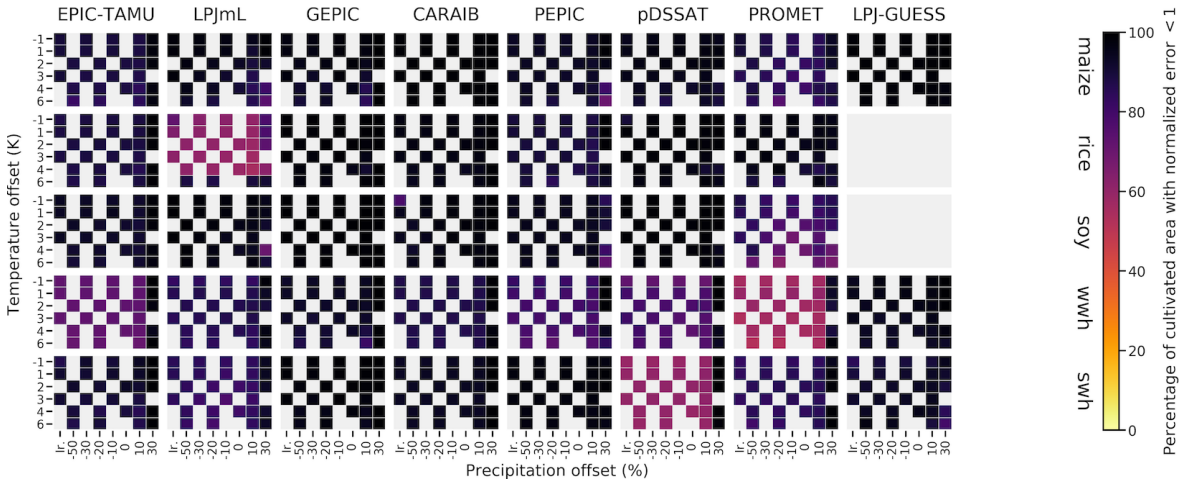


Figure S24: Fraction of currently cultivated hectares with normalized emulation error less than 1 for A1 yield emulation for [CO_2]=310 ppm and 200 kg N ha^{-1} yr^{-1} case. Figure convention as in Figure 3.7.

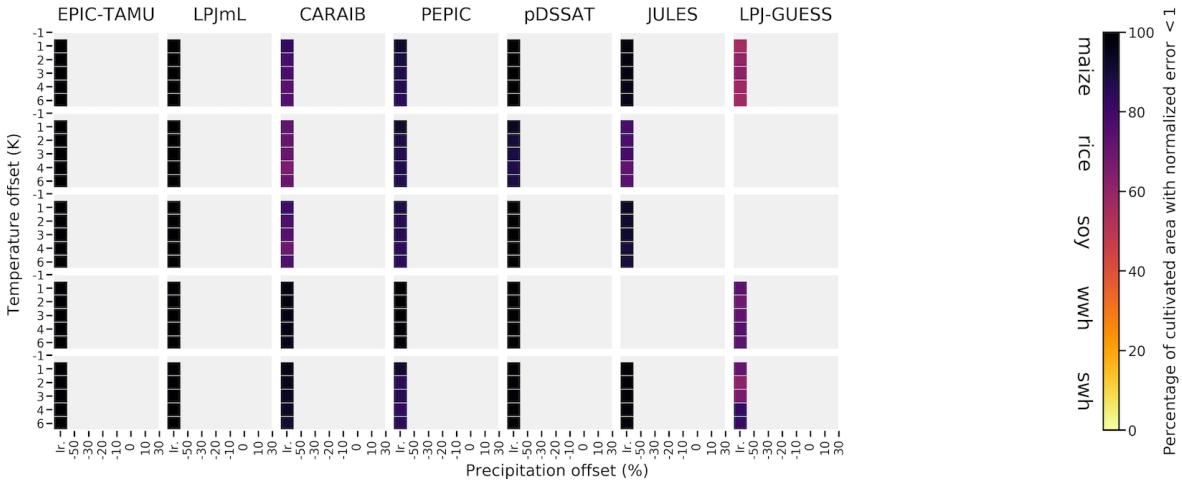


Figure S25: Fraction of currently cultivated hectares with normalized emulation error less than 1 for irrigated water demand emulation for [CO_2]=310 ppm and 200 kg N ha^{-1} yr^{-1} case. Figure convention as in Figure 3.7.

S2.5 Emulation of Yields in a Realistic Climate Simulation at High Latitude

In Chapter 3, we test the emulator against crop model simulations driven by a more realistic future climate projection to evaluate the impact of future variability changes that are not captured by the emulator. Figure S26 below isolates the mid- and high latitudes; compare to Figure 3.9 that shows global currently cultivated land. Results are generally unchanged by the restriction in latitude except for rice, which is typically grown in tropics and subtropics: only 20% of global rice production is grown north of 30N and 1% north of 45N, with even less in the Southern hemisphere, only 0.8% south of 30S and none south of 45S.

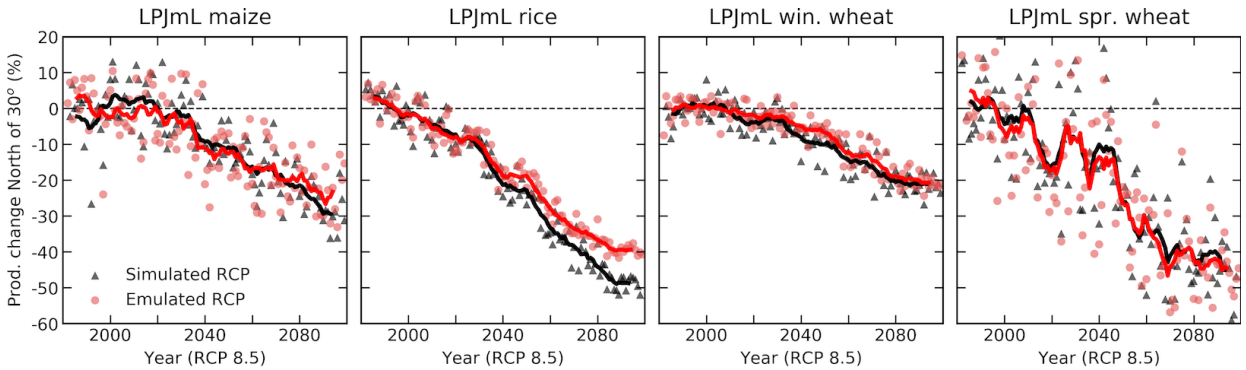


Figure S26: Illustration of the ability of the emulator to capture a more realistic future climate simulation, as in Figure 3.9 but here restricted to latitudes north of 30N.

S2.6 Emulator Products

This section amplifies the Chapter 3 with additional figures analogous to Figures 3.10 and 3.11.

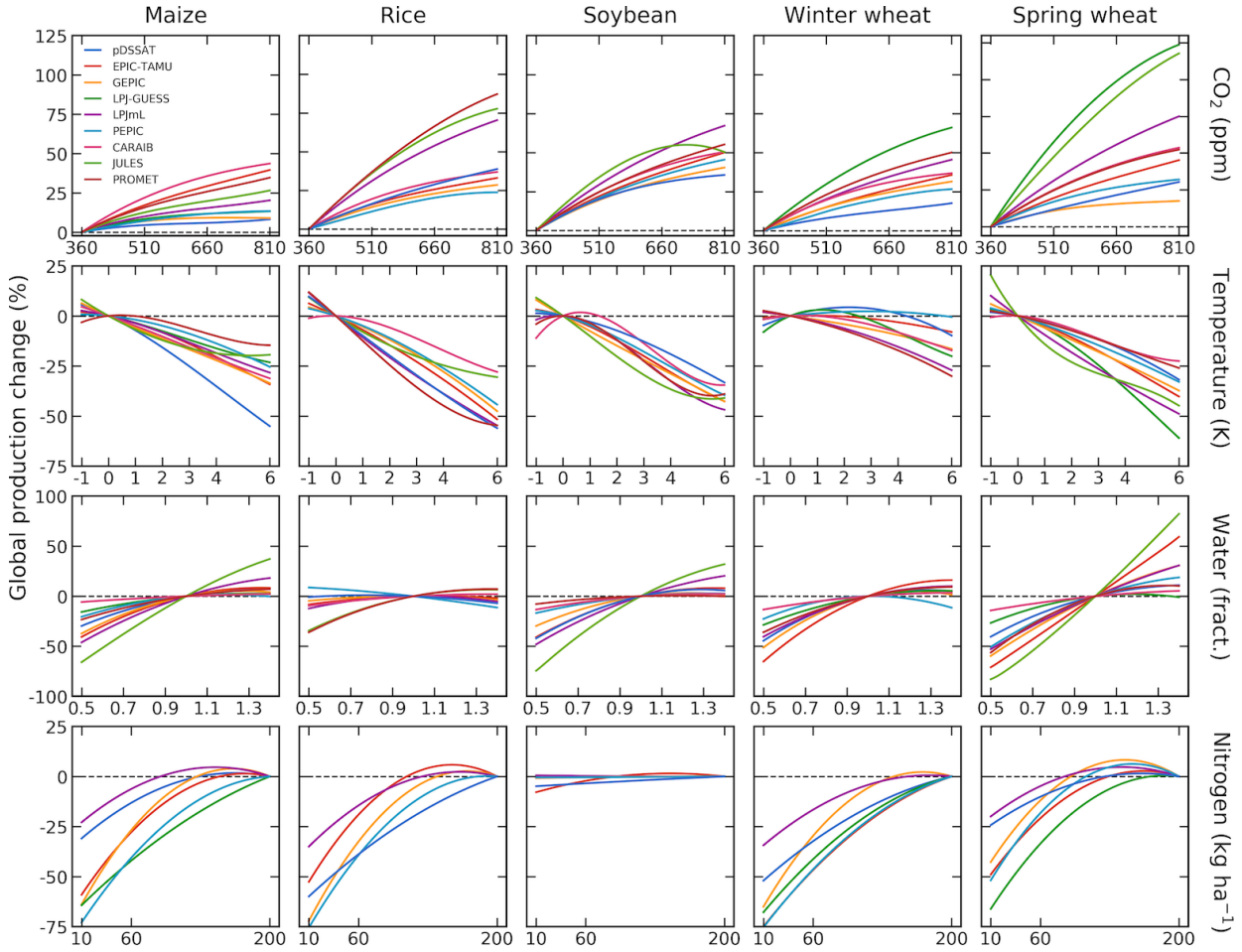


Figure S27: As in Figure 3.10 except now each model is shown in color. Not all models simulate every crop (JULES does not simulate winter wheat, and LPJ-GUESS does not simulate rice or soy) and three do not simulate the N response (CARAIB, JULES, and PROMET). For crops simulated, JULES and LPJ-GUES are often the outliers in strong [CO_2] responses. LPJmL has the weakest N response for all crops, and pDSSAT has the strongest T response for maize.

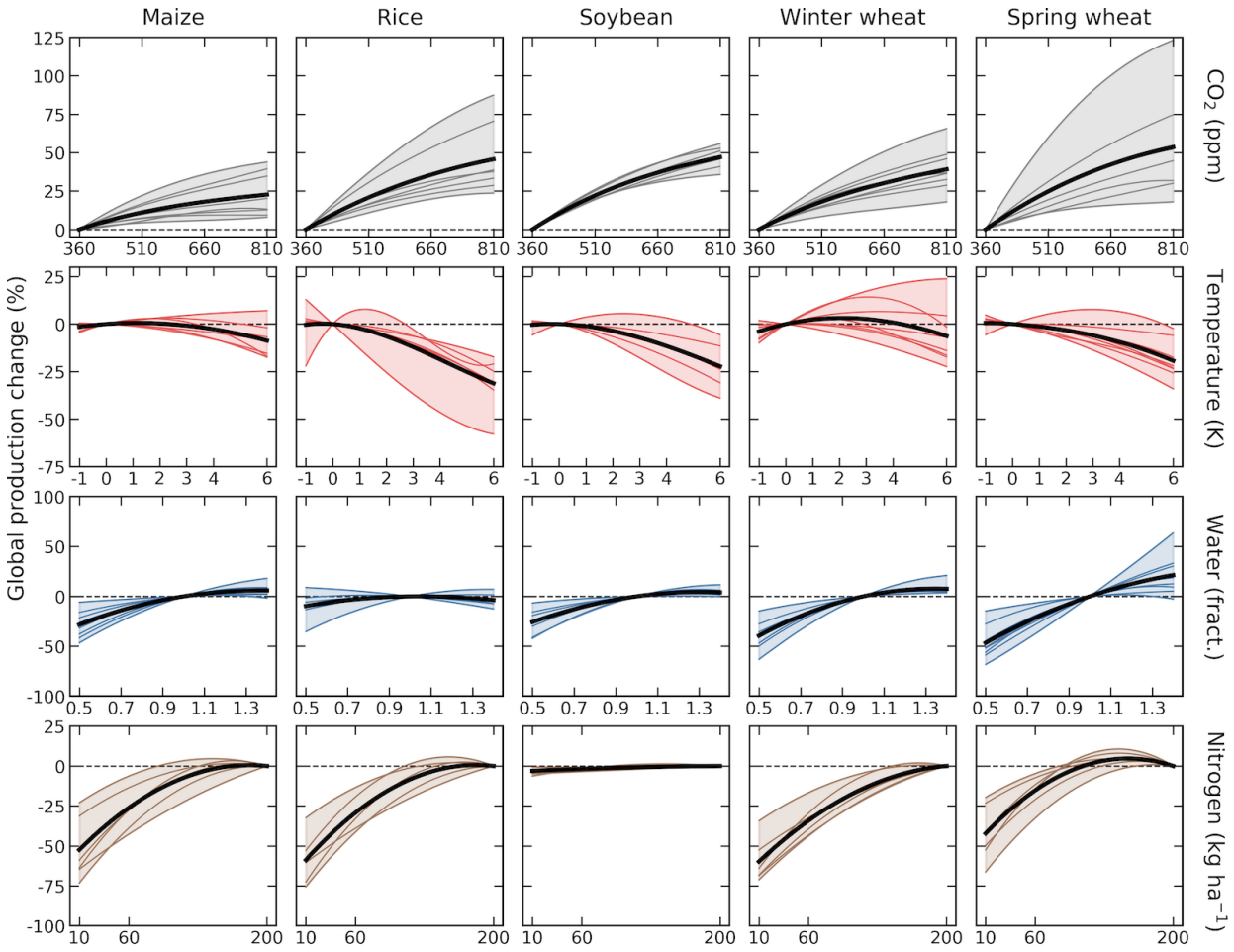


Figure S28: Emulated damage functions for rainfed crops for A1 (growing season adaptation) simulations, with conventions as in Figure 3.10 showing A0. Temperature responses are generally flatter than for A0 simulations, but responses to other factors are similar. Note that JULES does not provide A1 simulations.

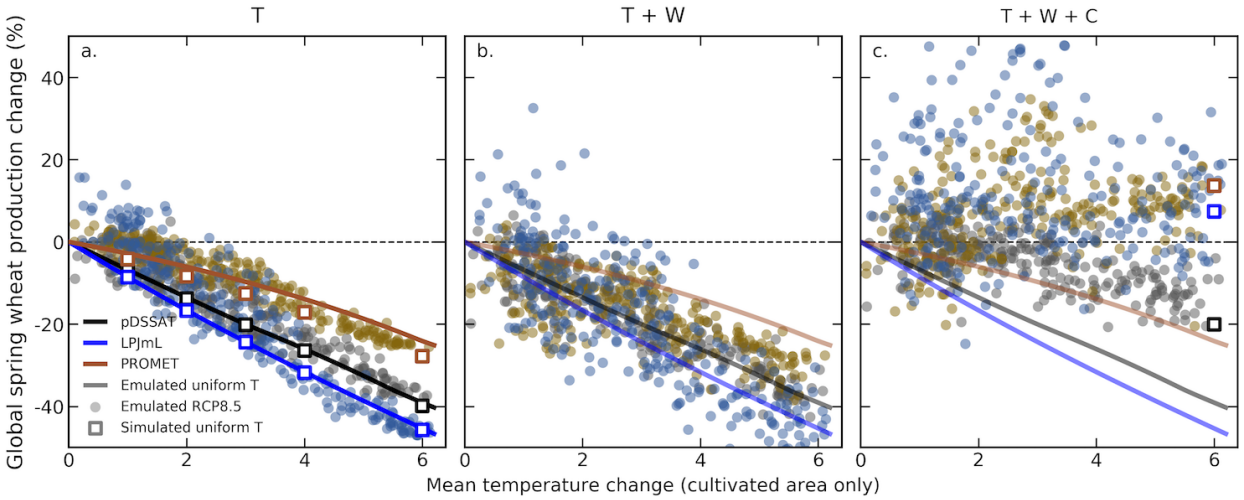


Figure S29: Illustration of the factors affecting yields in more realistic climate scenarios for rainfed and irrigated (current mix) spring wheat. Conventions as in Figure 3.11. Large emulator errors in PROMET spring wheat temperature response (panel **a**, compare open squares to line) are driven by Southern China, where discontinuities in yield responses make emulation problematic.

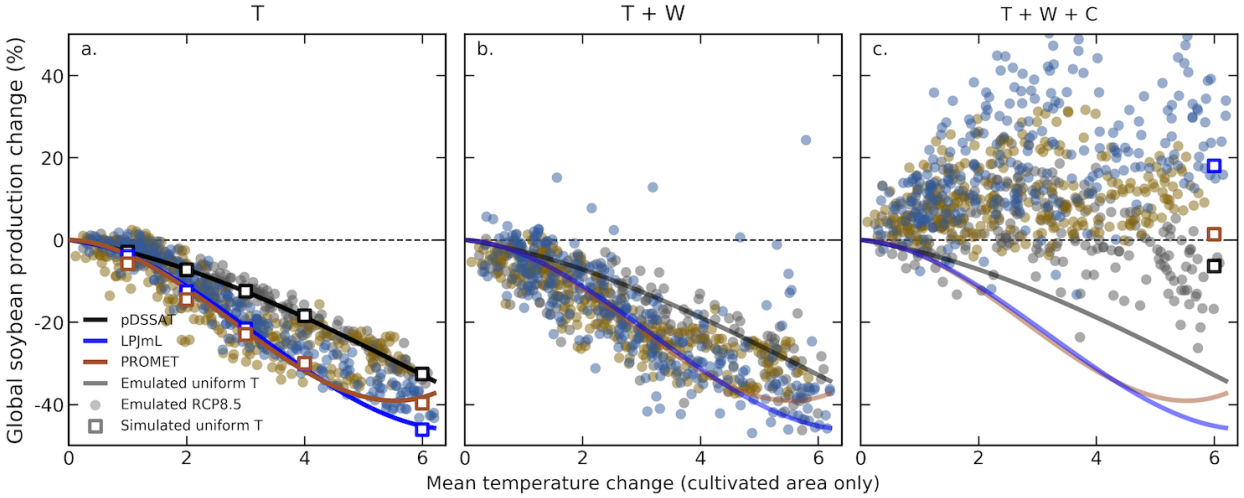


Figure S30: Illustration of the factors affecting yields in more realistic climate scenarios for rainfed and irrigated (current mix) soy. Conventions as in Figure 3.11. The split in PROMET soybean temperature response (panel **a**, note distinct groups of points) results from the model's sensitivity to differences in spatial patterns of temperature change across climate models.

S2.7 Reduced Specification (23-term) Emulator Examples

In this section we present analogous figures to those in the Chapter 3 for the reduced-form (23-term) emulator. Issues with the reduced-form model are most prominent in PROMET for rice and soy, and JULES soy and spring wheat. We identify several potential factors that may in some way contribute to these models showing qualitatively different responses that require additional terms for emulation.

- PROMET and JULES do not allow nitrogen variation. (However, CARAIB also cannot vary N and is readily emulatable with the 23-term specification.)

- Both JULES and PROMET models are land system process models, originally developed with a broad focus, which have been adapted for managed vegetation (agriculture) only recently (2015). (CARAIB, by contrast, was originally developed as a vegetation model in the early 90's and has a longer history of agricultural focus.)

- Both PROMET and JULES have anomalously strong responses to individual factors in those crops problematic to emulate. PROMET is the most sensitive model of all the models for rice in C, T, and W, and JULES for soybeans in C, T, and W. For spring wheat, JULES is a high outlier in C, the most sensitive model in W and T, and shows an extra inflection point in the global temperature response not seen in any of the other models.

- PROMET is the quantitatively lowest-performing model for soybeans when compared to the historical FAO data for the top 10 producing countries.

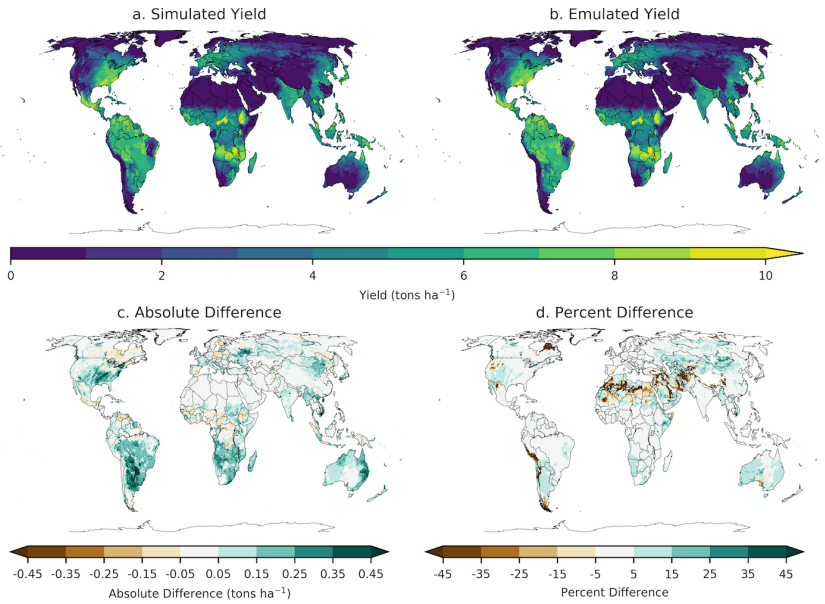


Figure S31: As in Figure 3.4, simulated (a.) and emulated (b.) yield under historical conditions for rainfed LPJmL maize, but here for the reduced (23-term) emulator specification. Emulator performance is worse primarily where crops are not currently grown.

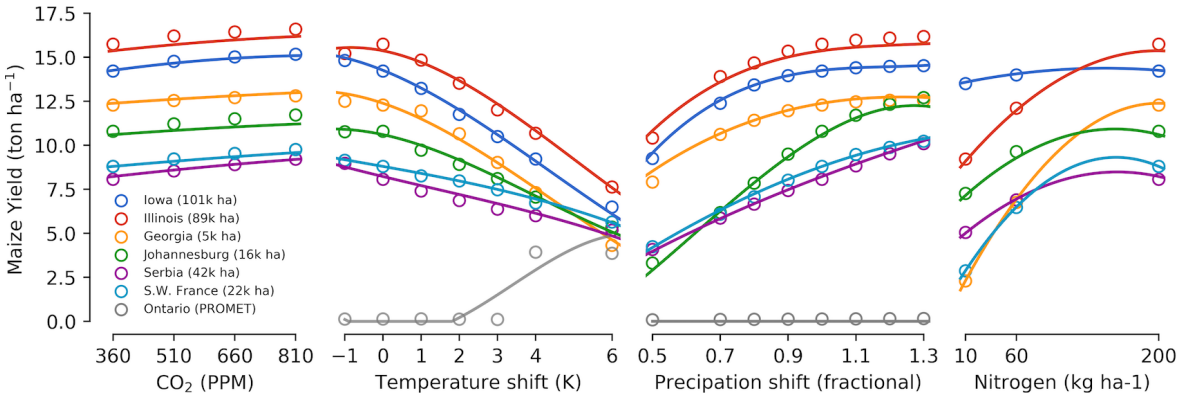


Figure S32: As in Figure 3.5, emulator performance in selected high-yield regions for rainfed pDSSAT maize (and one region for PROMET), but now with the reduced (23-term) emulator specification. Emulator performance is similar.

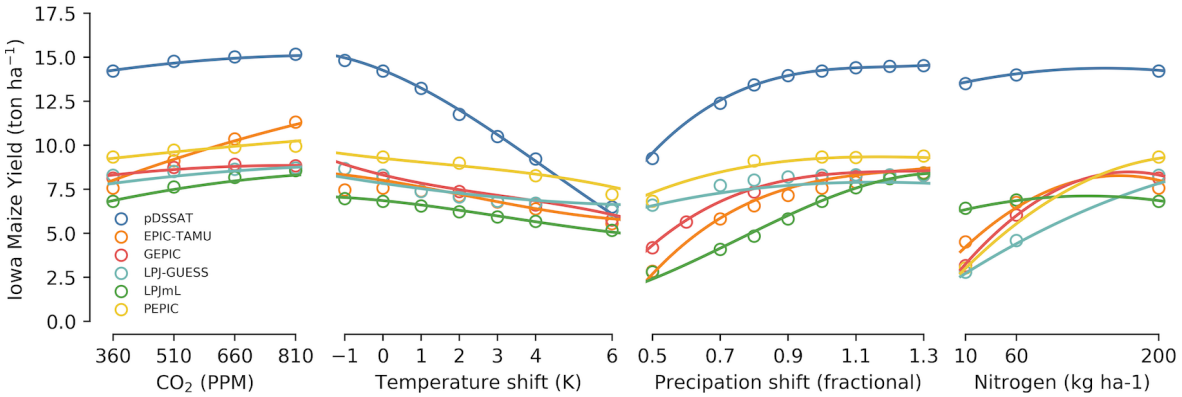


Figure S33: As in Figure 3.6, emulator performance across models for rainfed maize in one grid cell in Iowa, but now with the 23-term emulator specification. Note that JULES and PROMET are not shown.

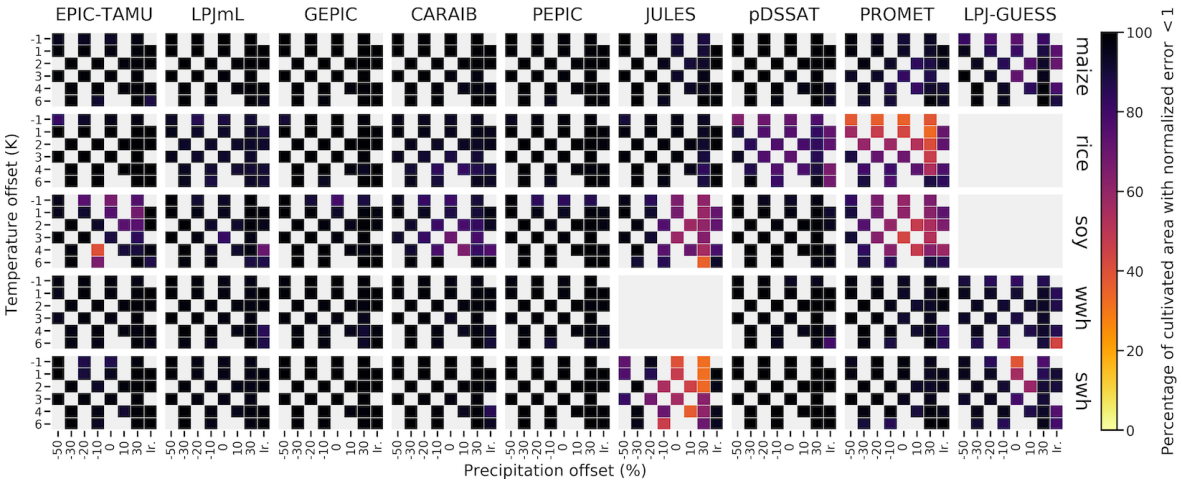


Figure S34: As in Figure 3.7, normalized error of all 9 models emulated on currently cultivated land, over all crops and all sampled T and W inputs, with [CO_2] and nitrogen held fixed at baseline values, now with the reduced (23-term) emulator specification. Degradation of performance is most evident in JULES soy and spring wheat and PROMET rice and soy.

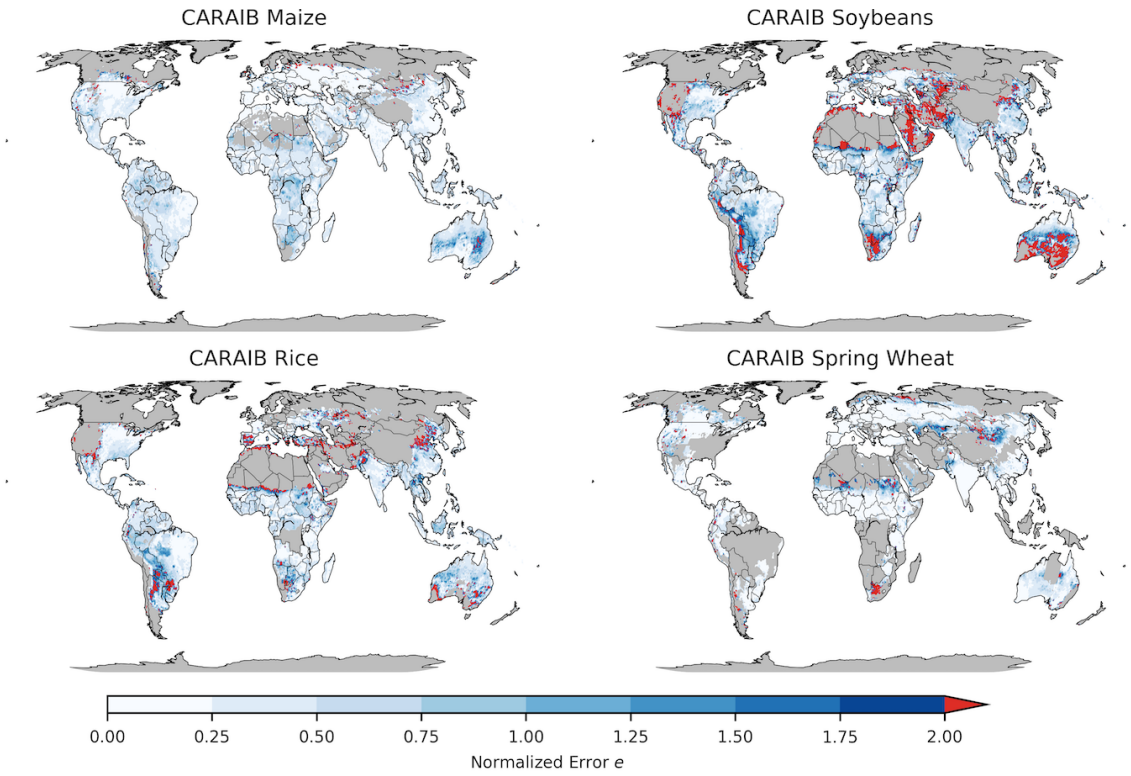


Figure S35: As in Figure 3.8, normalized error for rainfed crops in CARAIB for the T+4 scenario, but here with the reduced (23-term) emulator specification. Degradation of performance is most evident in marginal lands where crops are not currently grown.

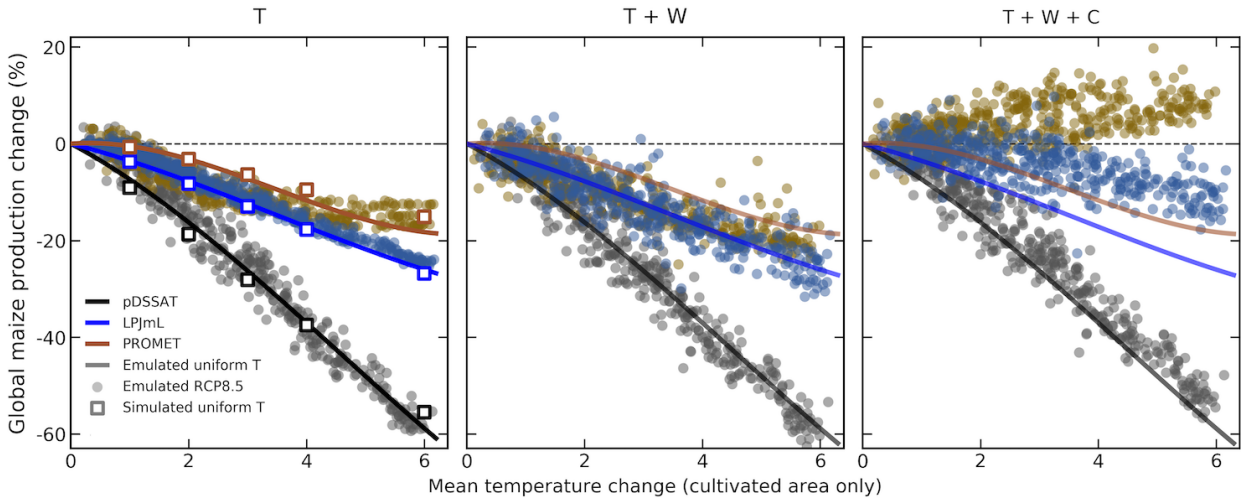


Figure S36: As in Figure 3.11, rainfed maize on currently cultivated land, but here with the reduced (23-term) emulator specification. Note that strong C response for PROMET is different here than with the full-form emulator, because higher order C (C^3, $C^2 * T$ ...) interaction terms are needed for accurate emulation.

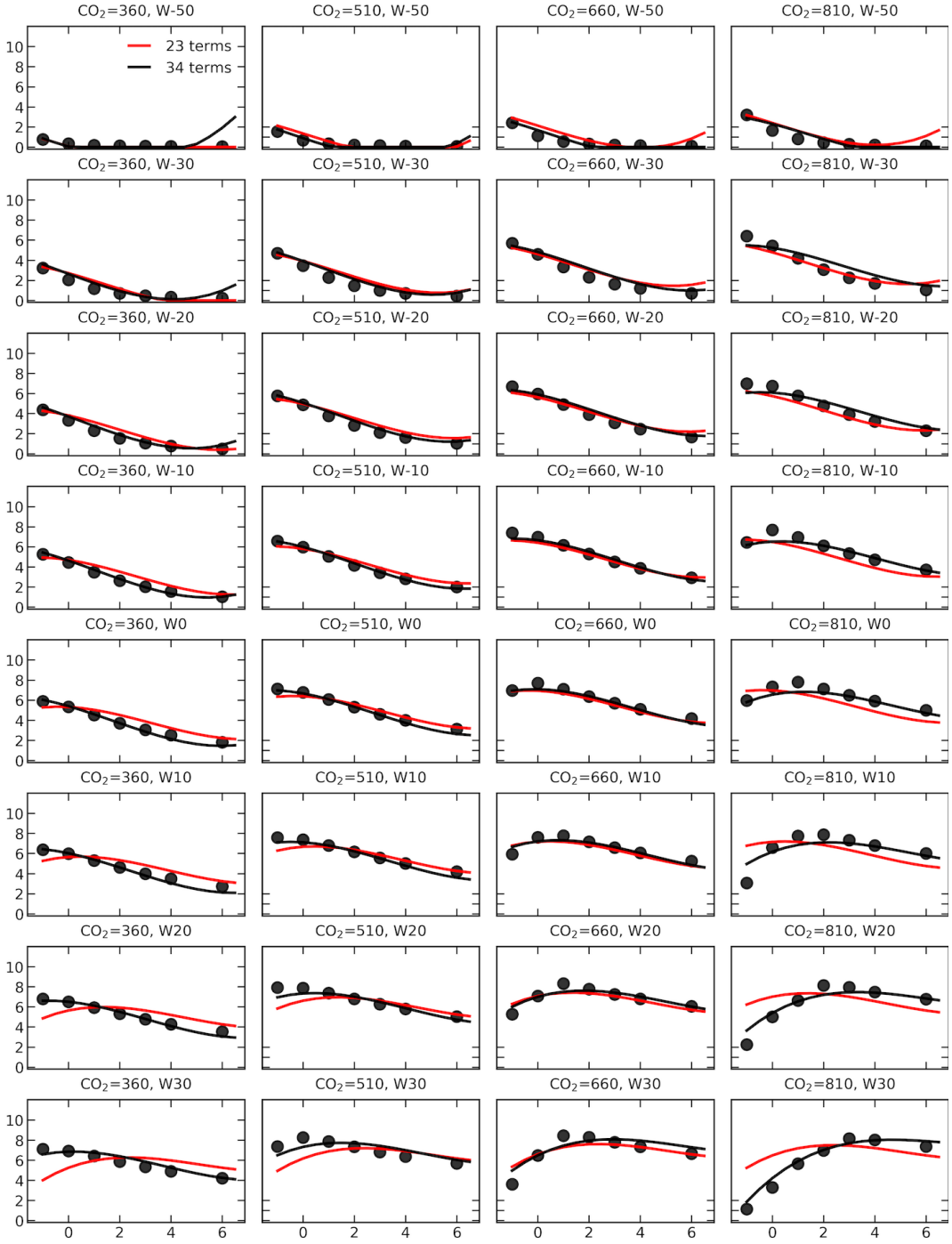


Figure S37: Example of emulator failure, showing failure induced because of strongly interacting terms. Simulated and emulated values for JULES soybeans in Southern Germany. RMSE = 41% of baseline yield for the reduced form (23-term) emulator. The downturn in yields as C and W increase can only be captured by the higher order C interaction terms of the full 34-term specification.

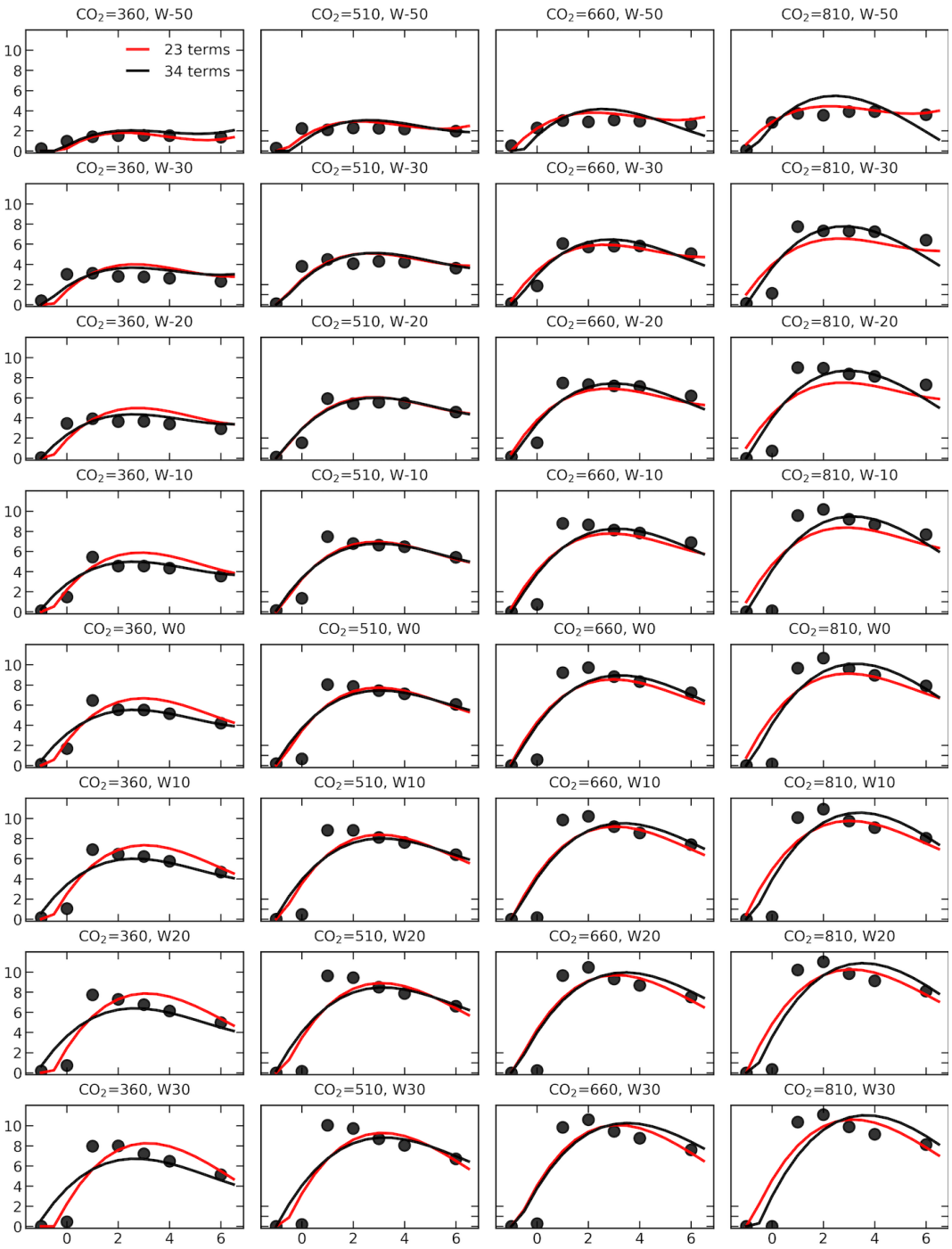


Figure S38: Example of emulator failure, showing failure induced by abrupt changes in yields. Simulated and emulated values for PROMET rice in India (Arunachal Pradesh). RMSE = 132% of baseline yield for red (reduced fit). The step change in the yields around 0 K at higher water specifications cannot be captured by any third order polynomial. Both 23- and 34- emulator specifications fail in this example.

S2.8 Yield Responses for Other Crops and Models

Spatial patterns of yields are well captured for all crops and models. Figure 3.4 illustrated this using LPJmL maize; for reference, we show here yield response spatial patterns for other crops and models.

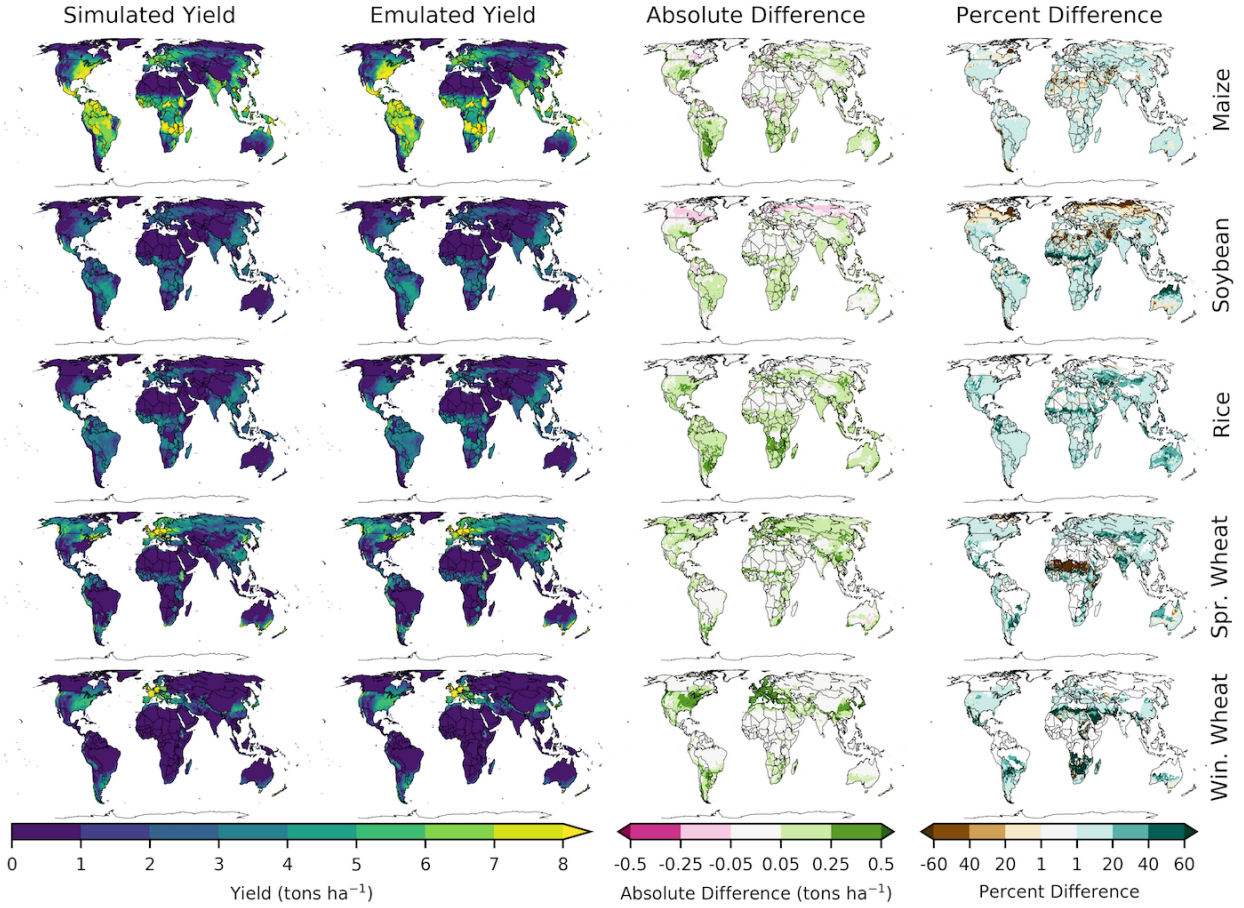


Figure S39: Spatial yield response and emulator error for LPJmL for all 5 GGCMI Phase II crops. Convention as in Figure 3.4.

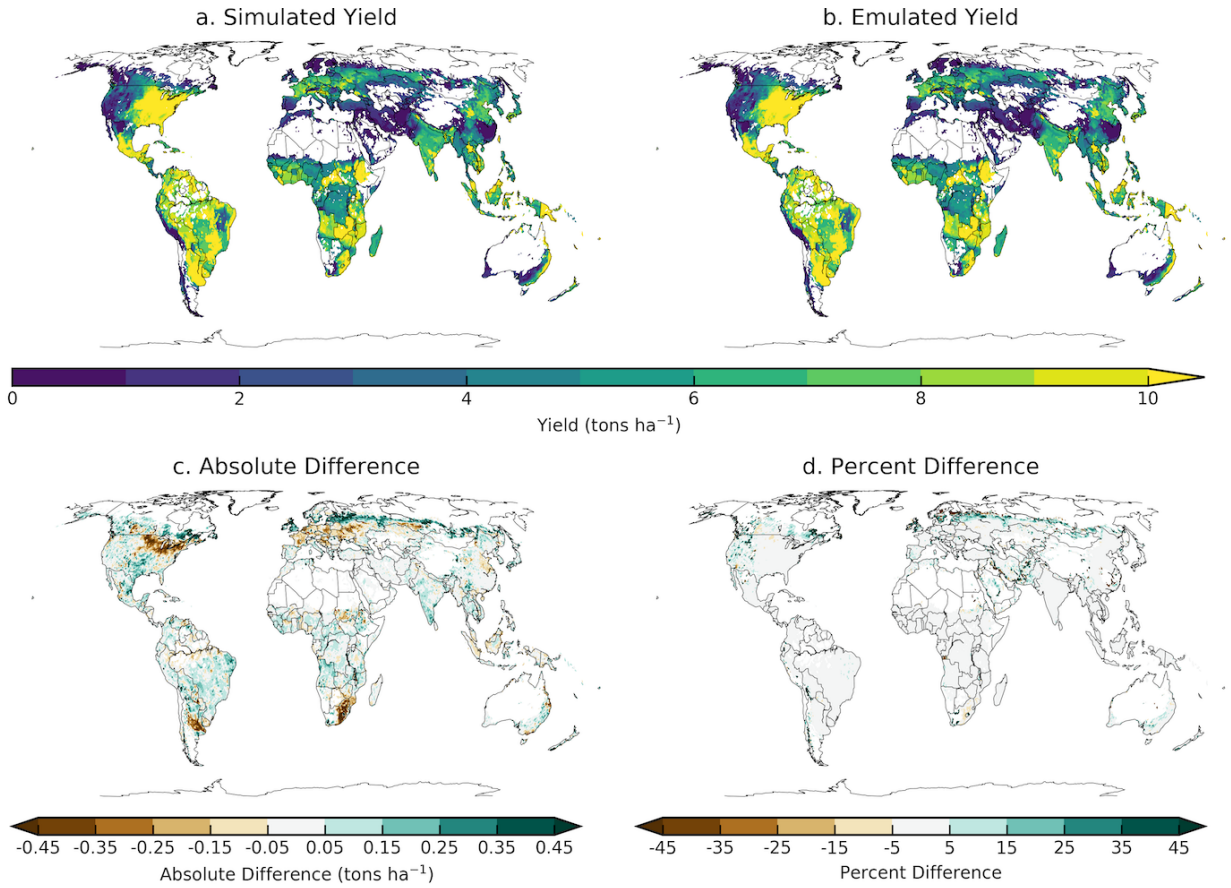


Figure S40: Spatial yield response and emulator error for pDSSAT for maize. Convention as in Figure 3.4. pDSSAT absolute yields are significantly higher than those in LPJmL but spatial patterns are similar.

S2.9 Cross Validation Error for All Models

In this section we present maps of cross validation error (values found in Table 3.1 are aggregated up from the grid cell level). Errors are generally low as a percentage of yield change in each grid cell. Errors above 10% of yield change in the out-of-sample test occur very rarely; the only significant instance is spring wheat in southern China in the PROMET model.

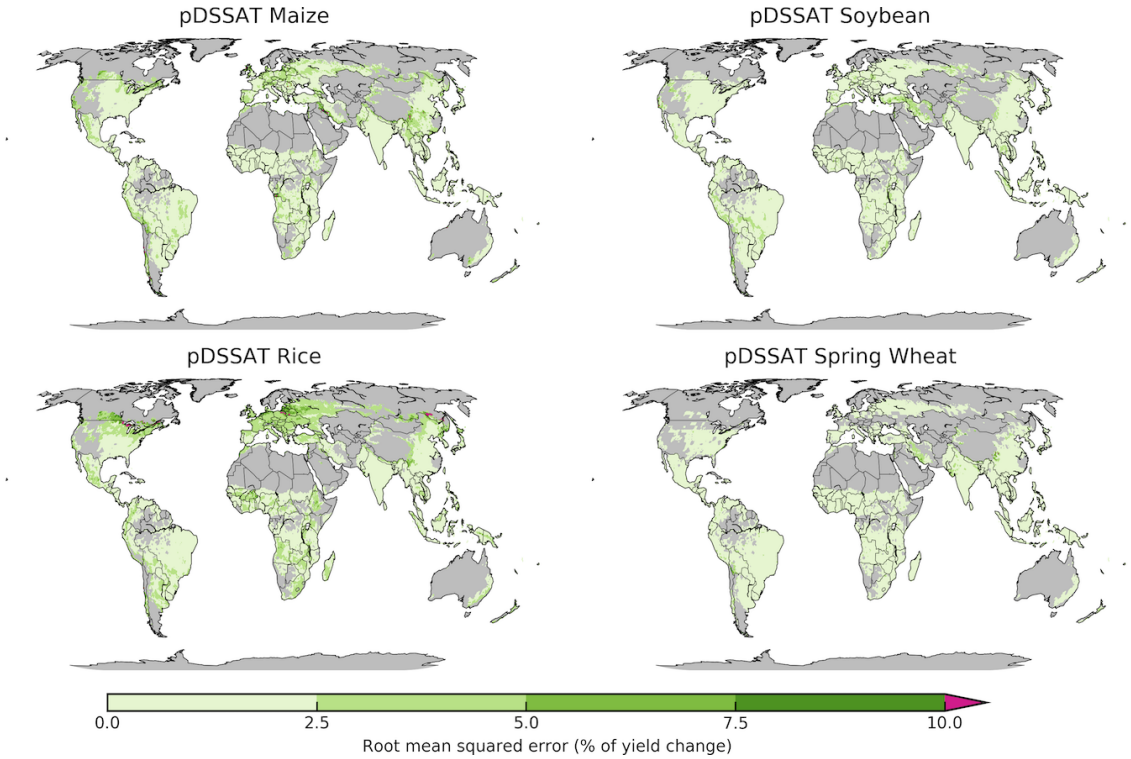


Figure S41: Root mean squared error for cross validation for the pDSSAT model for rainfed crops. Values shown as a percentage of yield change in each grid cell.

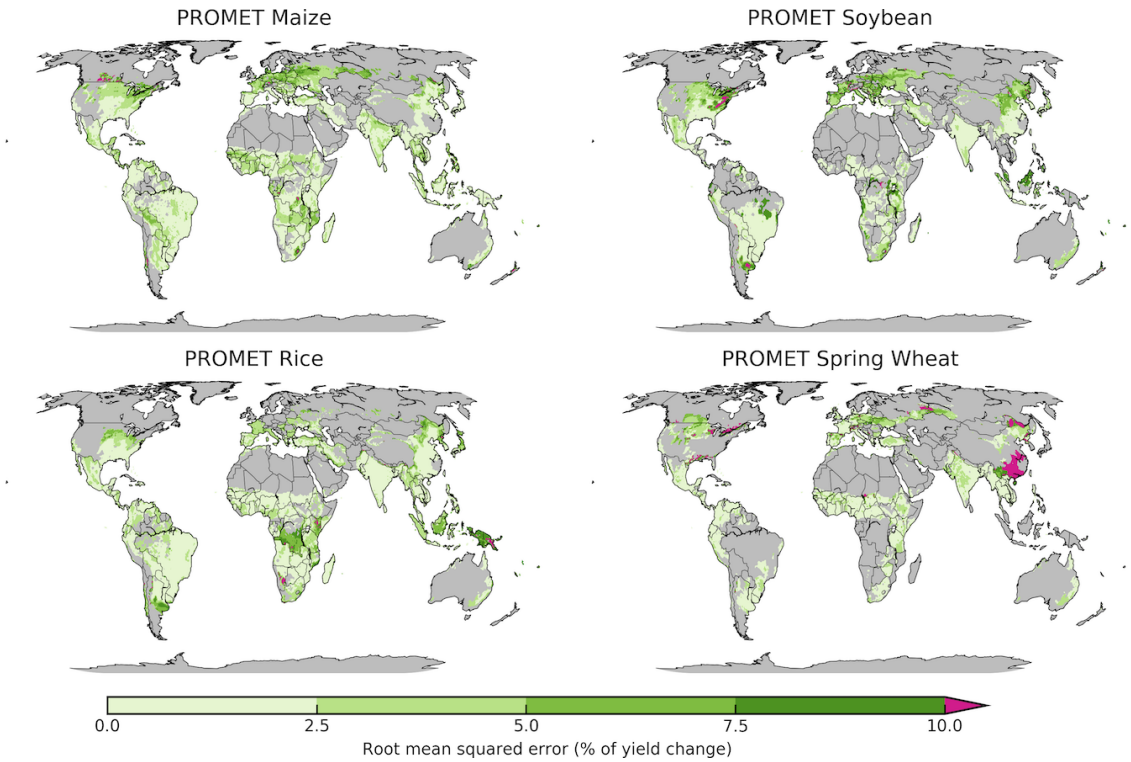


Figure S42: Map of root mean squared error for cross validation process for the PROMET model for rainfed crops. Values shown as a percentage of yield change in each grid cell.

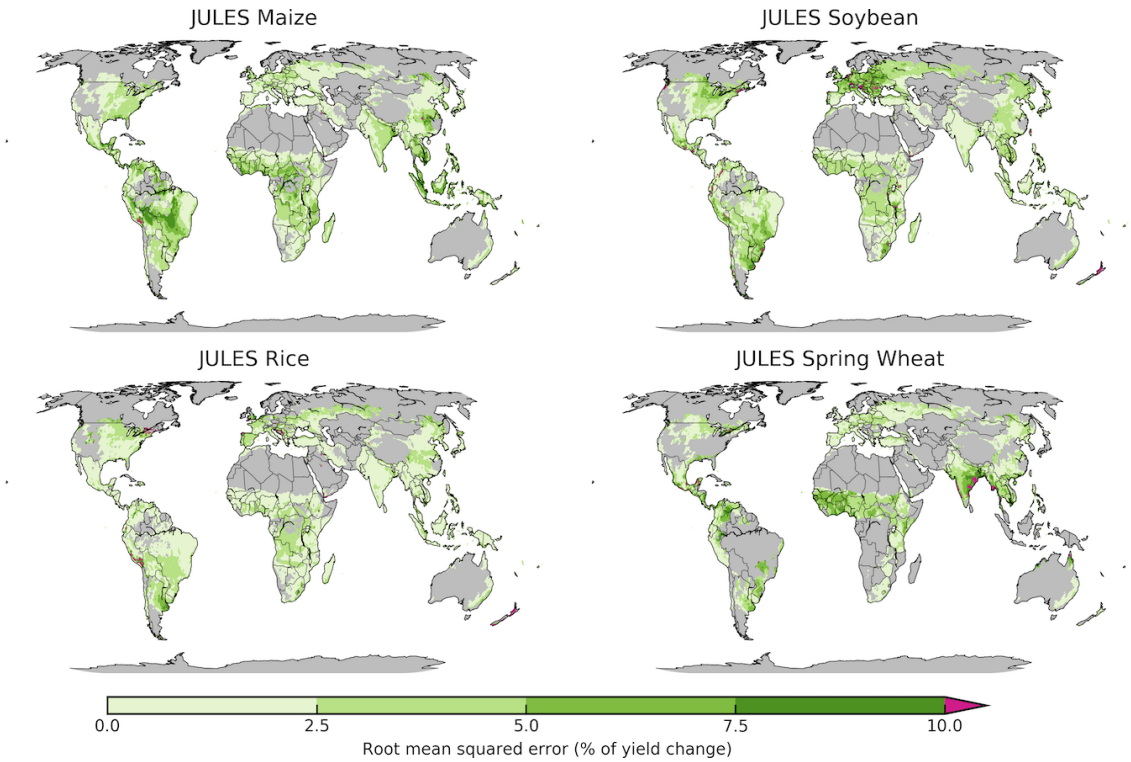


Figure S43: Map of root mean squared error for cross validation process for the JULES model for rainfed crops. Values shown as a percentage of yield change in each grid cell.

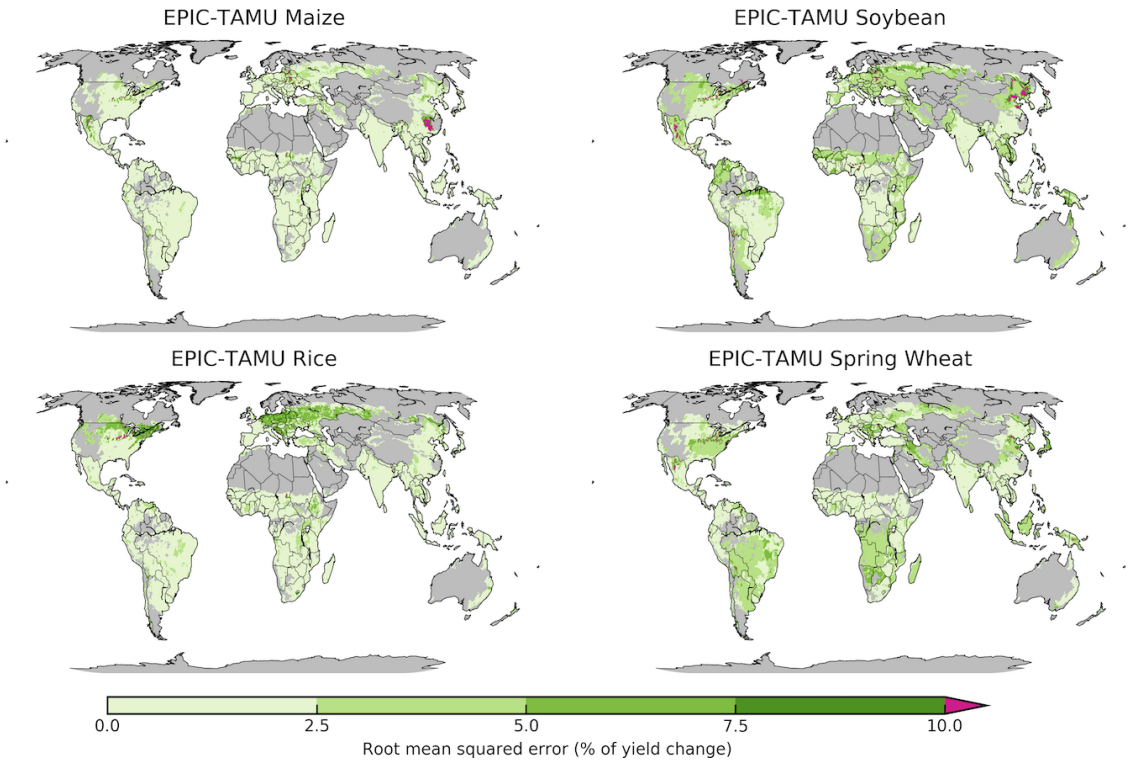


Figure S44: Map of root mean squared error for cross validation process for the EPIC-TAMU model for rainfed crops. Values shown as a percentage of yield change in each grid cell.

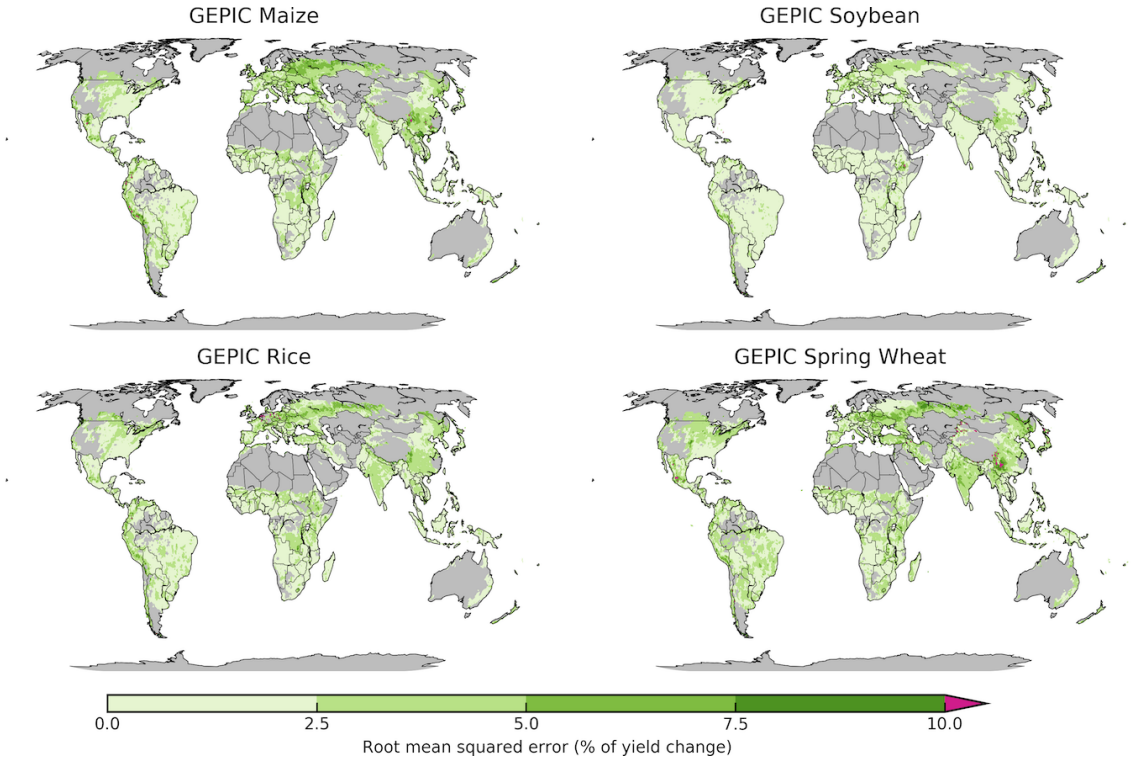


Figure S45: Map of root mean squared error for cross validation process for the GEPIC model for rainfed crops. Values shown as a percentage of yield change in each grid cell.

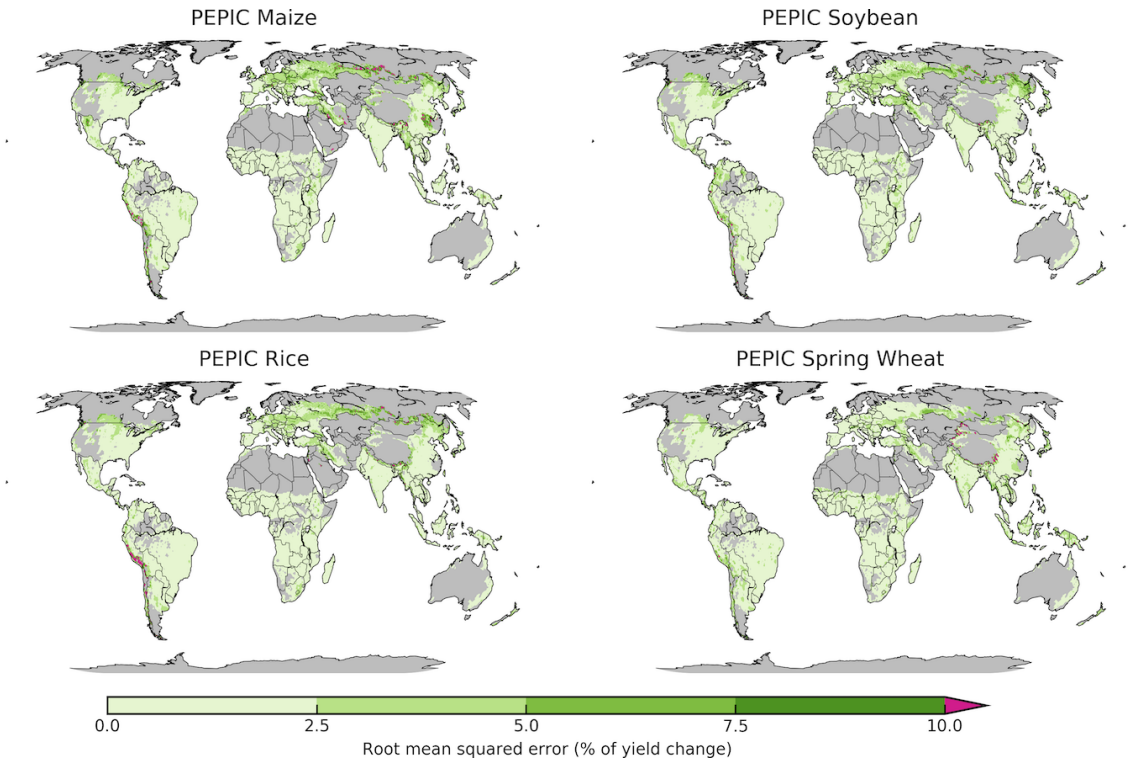


Figure S46: Map of root mean squared error for cross validation process for the PEPIC model for rainfed crops. Values shown as a percentage of yield change in each grid cell.

S3 Supplemental Materials for Chapter 4

S3.1 Additional LPJmL Simulations

In addition to the GGCMI Phase 2 data set, we ran the LPJmL model for additional scenarios to allow for better analysis of the mechanisms that determine future spatial yield patterns under climate change. One important shortcoming of the GGCMI Phase 2 data set is that the only adaptation mechanism considered is the regaining of the growing season that shortens under warming due to accelerated phenological development (A1), while the adaptation in sowing dates is ignored in this data set. Also, regaining the growing season under warming must not necessarily be a useful adaptation measure, as shorter growing seasons could also be beneficial, e.g. to avoid terminal drought.

For the extra simulations, we do not reproduce the GGCMI experimental setup with regular disturbances of the historical AgMERRA climate data set along the T and W dimensions, but conduct simulations for transient climate projections from the ISIMIP2b project, making use of bias corrected climate scenarios of HadGEM, MIROC, GFDL, and IPSL for the high-end radiative forcing scenario RCP8.5. For each of these climate scenarios, we developed adapted planting date and harvest date scenarios for the end of the 21st century, making use of the rule-based approach. These scenarios assume no delay in cultivar development, so that optimal growing seasons can be exploited with suitably fast/slow maturing cultivars. Simulations were conducted for the period 1980-2099, covering the historical reference period (1980-2010) and the future period (2070-2099). For the historical period, we use rule-based rather than observed growing seasons (as done in the GGCMI Phase 2 experiment), to avoid comparing apples and oranges when comparing historical and future yield patterns. We conducted simulations where only the sowing dates (planting advance), only the cultivar (maturity delay) or both (plant & maturity) were adapted, in order to allow for separating the effects.

On top of these growing season adaptation scenarios, we conducted simulations with modified input data, where all gridcells were parameterized to have the soil properties as well as the same radiation inputs of Cedar Rapids, IA, USA. With this unrealistic simulation setup (soils rad.), we intended to see if the remaining differences of yield levels with adapted growing seasons and shifted breadbasket locations could be explained by centrally important, but non-transferable properties. This setup is computed for non-adapted growing seasons as well as in combination with the full growing season adaptation (plant & maturity).

Table S3: Growing season changes for the LPJmL simulations driven by the HadGEM climate model. GSL = Growing season length, time between planting and maturity.

Simulation Case	Plant day (doy)	Harvest day (doy)	GSL (days)	Planting advance (% of GSL)	GSL change (%)	T change (K)	Yield change (%)	Poleward shift (km)
Historical (1985-2005)	122	244	122	NA	NA	NA	NA	NA
No adaptation (A0) (2080-2100)	122	206	84	0	-30	6.7	-20	55
Maturity delay (A1) (2080-2100)	122	240	118	0	-3	4.0	-15	610
Planting advance (2080-2100)	101	193	92	+16	-25	1.9	-15	400
Plant + maturity (2080-2100)	101	232	132	+16	+8	2.8	+3	780
Plant + maturity + soils + radiation (2080-2100)	101	232	132	+16	+8	2.3	+8	910

S3.2 Breadbasket and Cultivation Areas

S3.3 Model Comparison and CMIP6 Climate Velocity

S3.4 Results

Table S4: Figure 4.6 summary statistics. NA = no adaptation. FA = full adaptation. Shift in land under full adaptation is the percentage of new grid cells which are now in the top 10% of yields at the end of the century.

Breadbasket	Poleward shift, NA (km)	Poleward shift, FA (km)	Shift in land, FA (%)	Production change over current land, NA (%)	Production change FA + shift (%)
North American Maize	55	780	66	-20	+8
South American Soybeans	955	777	46	-23	+26
European Wheat	-290	344	35	-2	+23
East Asian Rice	-122	511	60	-25	+20

Table S5: Percent of area within the breadbasket on the continental scale as shown in Figure 4.6. Some cautions should be taken with interpreting these numbers. The breadbaskets in our analysis are defined as the highest-yielding decile of all grid cells (regionally) – not by their actual production area and we also do not project future shifts in land-use patterns. How much of a current breadbasket shifts across national borders depends not only on the strength of that shift, but also by the size and geometry of countries, as well as given distortions of yield patterns in GGCMs.

% of area shifting across national borders	No adaptation	Full adaptation
Maize (North america)	0	23
Soybean (South America)	80	67
Wheat (Western Europe)	47	62
Rice (Eastern Asia)	15	30

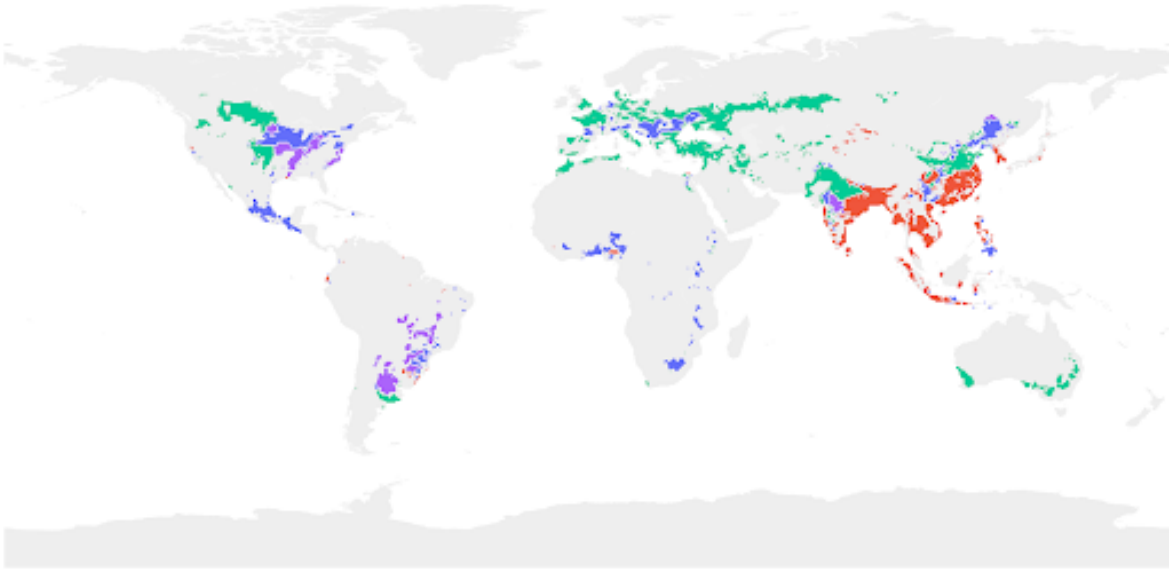


Figure S47: Major Breadbasket locations. Top 20% cultivation intensity regions globally for maize (blue), soybean (purple), wheat (green), and rice (red). Data from MIRCA2000.

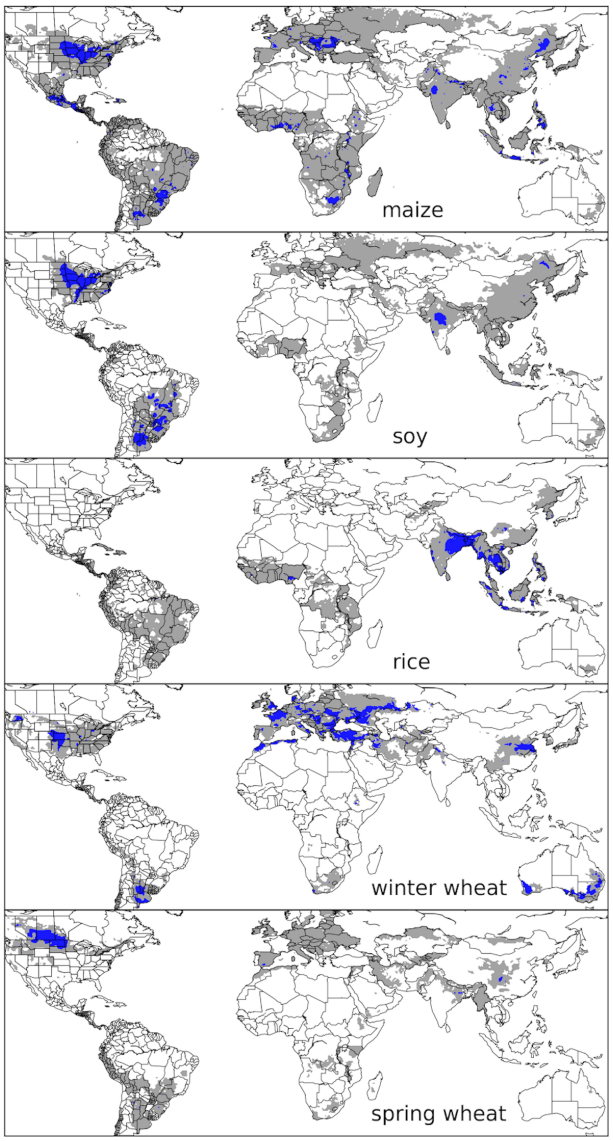


Figure S48: Presently cultivated area for rainfed crops from MIRCA2000. Blue indicates grid cells with more that 20,000 hectares (10% of the equatorial grid cell). Gray contour shows the area at least 10 hectares cultivated per grid cell. Data from MIRCA2000.

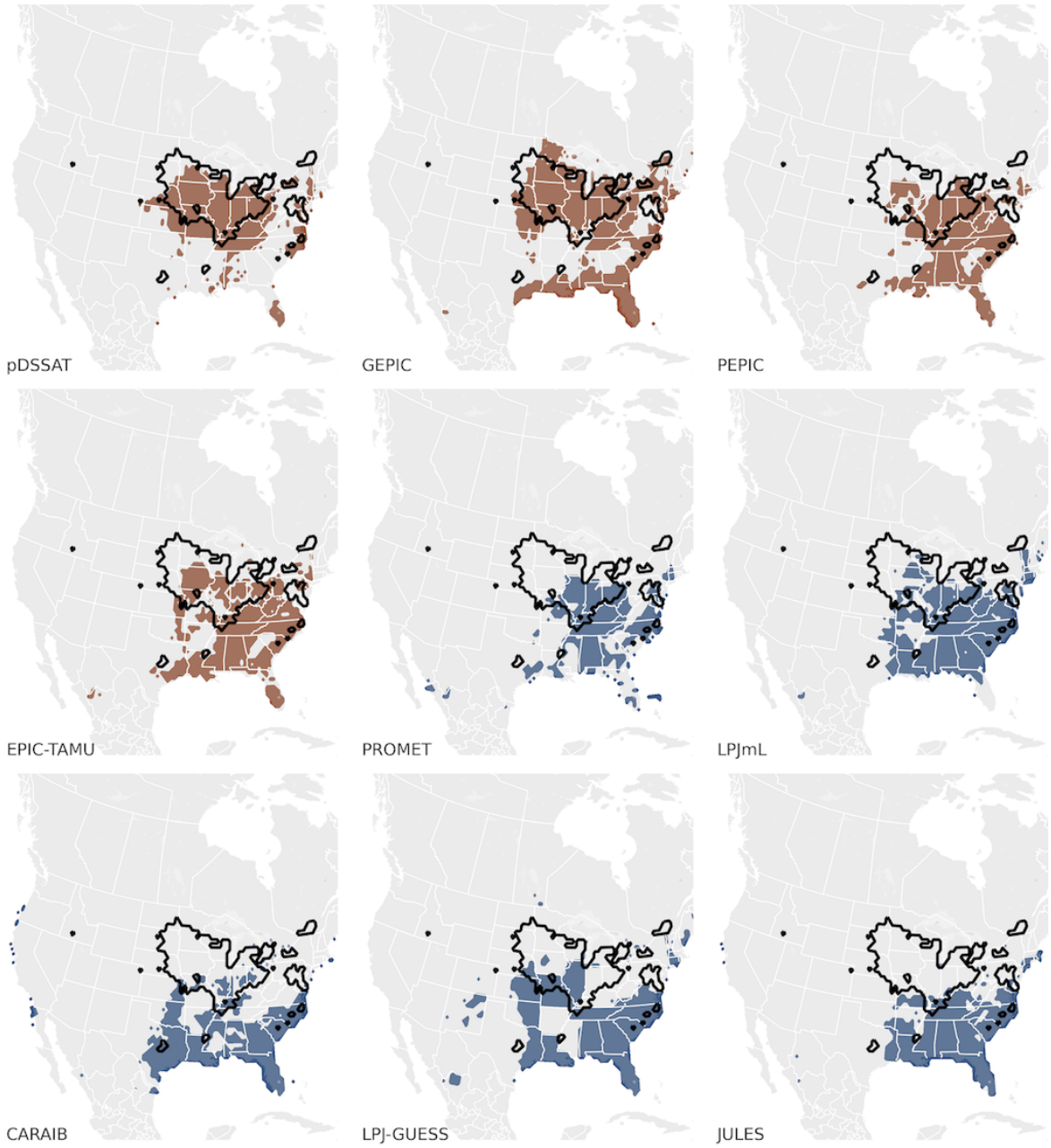


Figure S49: Ecosystem vs site based model comparison, historical simulations. Top 10% simulated grid cells in North America for each model for rainfed maize under historical conditions (1980-2010). Black contour is the top 15% rainfed cultivation intensity for reference. Site-based models generally do a better job of locating the best yields in the midwest corn belt, though all models other than pDSSAT place much of the best yields on the gulf coast.

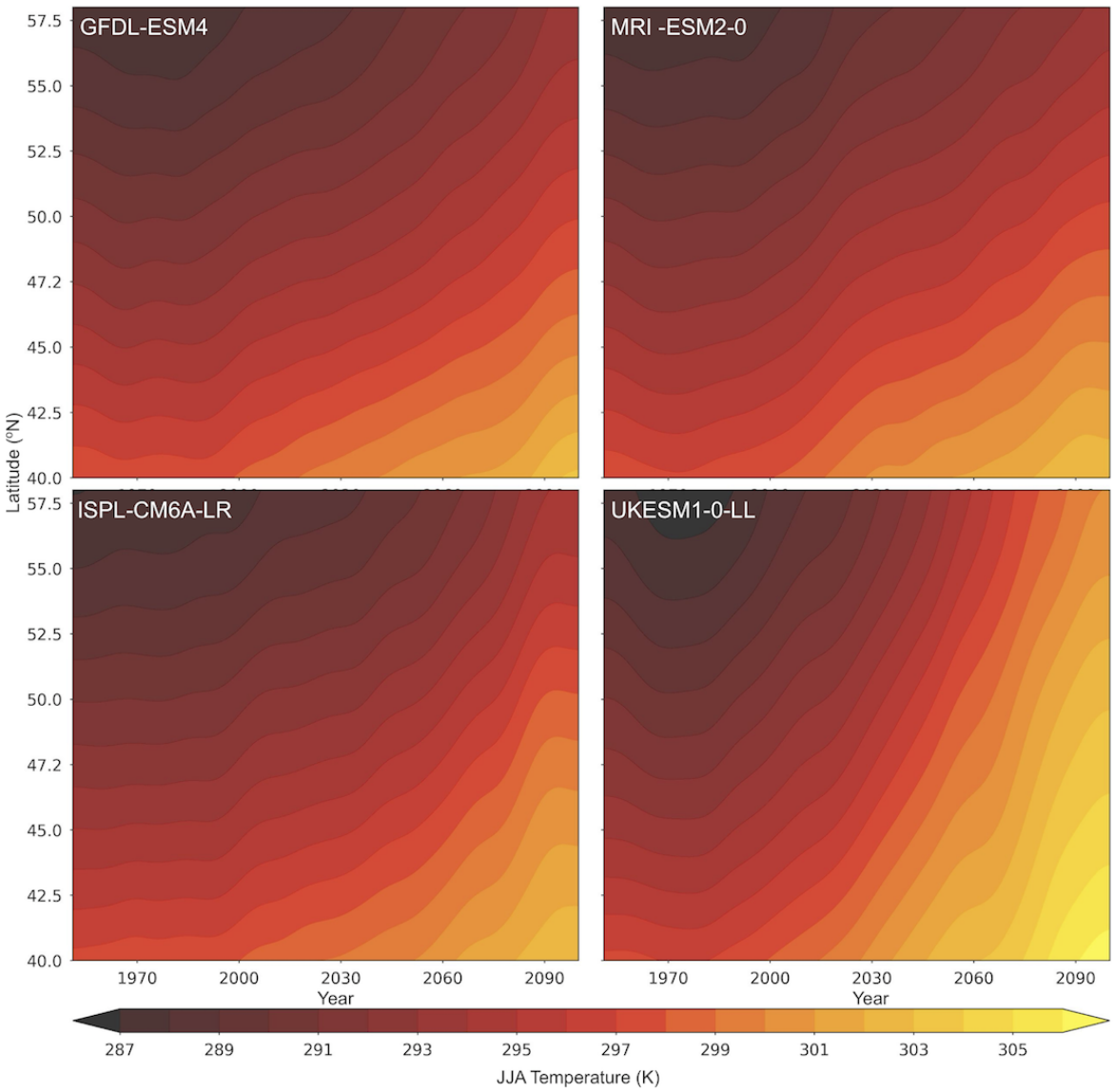


Figure S50: CMIP6 temperature shifts along latitudes (SSP585). June-July-August zonal mean temperature in North America (over only agricultural land) for select CMIP6 models for SSP585. Isotherms show poleward progression of growing season temperature analogue. Yearly temperatures smoothed with a gaussian filter to show climatology.

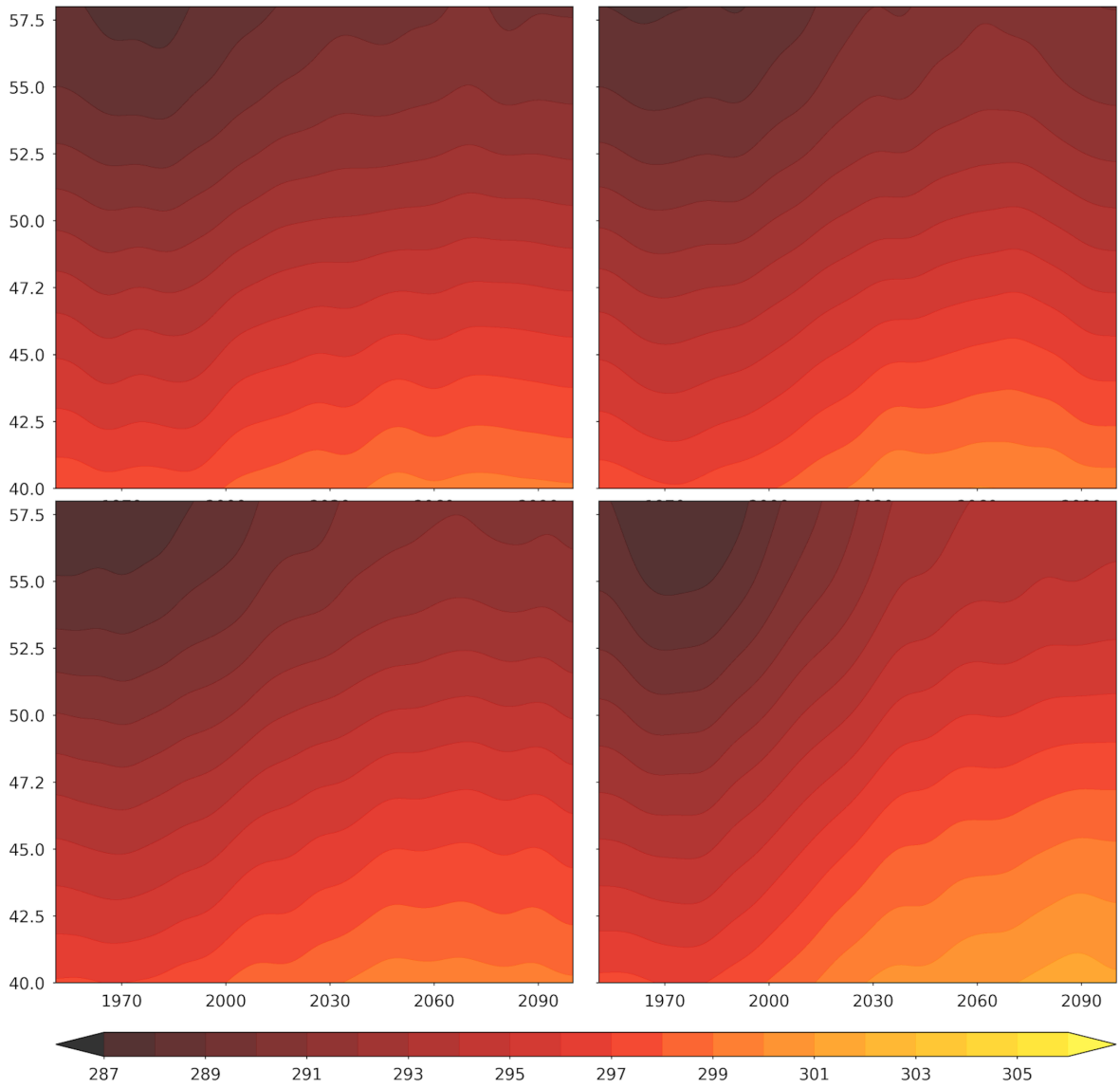


Figure S51: CMIP6 temperature shifts along latitudes (SSP126). Same as above, now for SSP126. UKESM (lower right quadrant, a high sensitivity model) shows 5oC of temperature rise at 42N (Cedar Rapids).

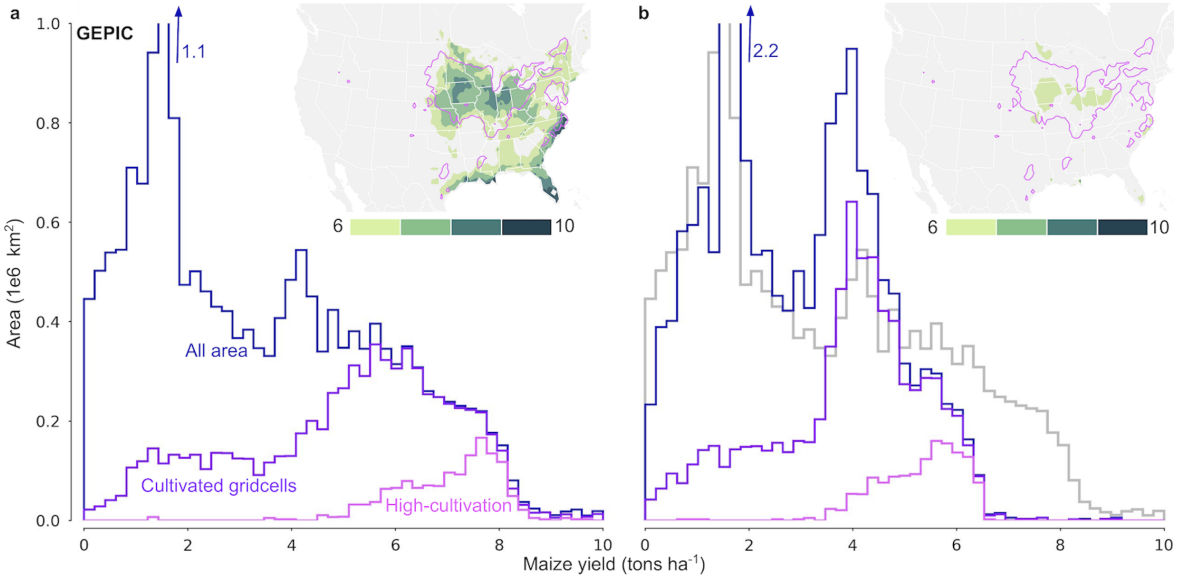


Figure S52: Same convention as Figure 4.2, except now for high [CO_2] of 810 ppm from the GEPIC model. Simulations are the A0 growing season specification.

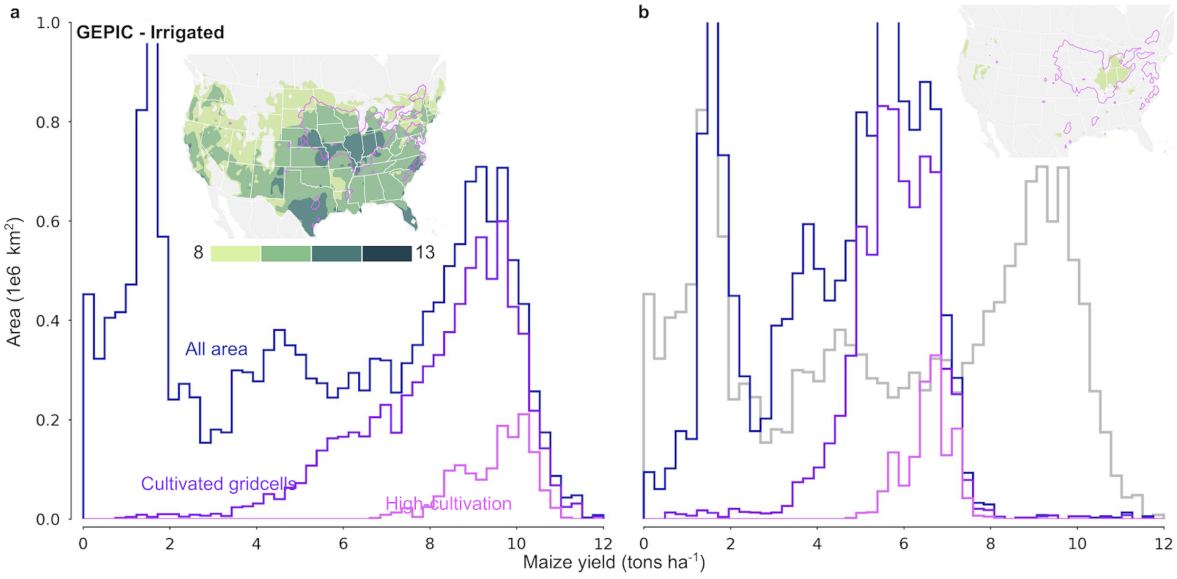


Figure S53: Same convention as Figure 4.2, except now for irrigated simulations from the GEPIC model. Simulations are the A0 growing season specification.

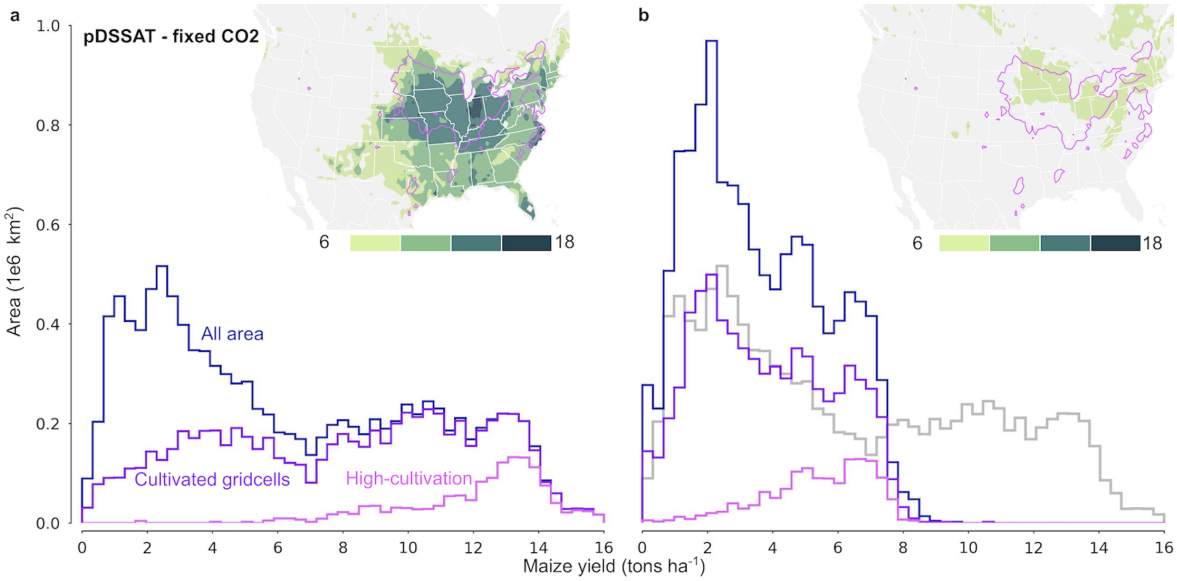


Figure S54: Same convention as Figure 4.2, except now for the pDSSAT model. Simulations are the A0 growing season specification.

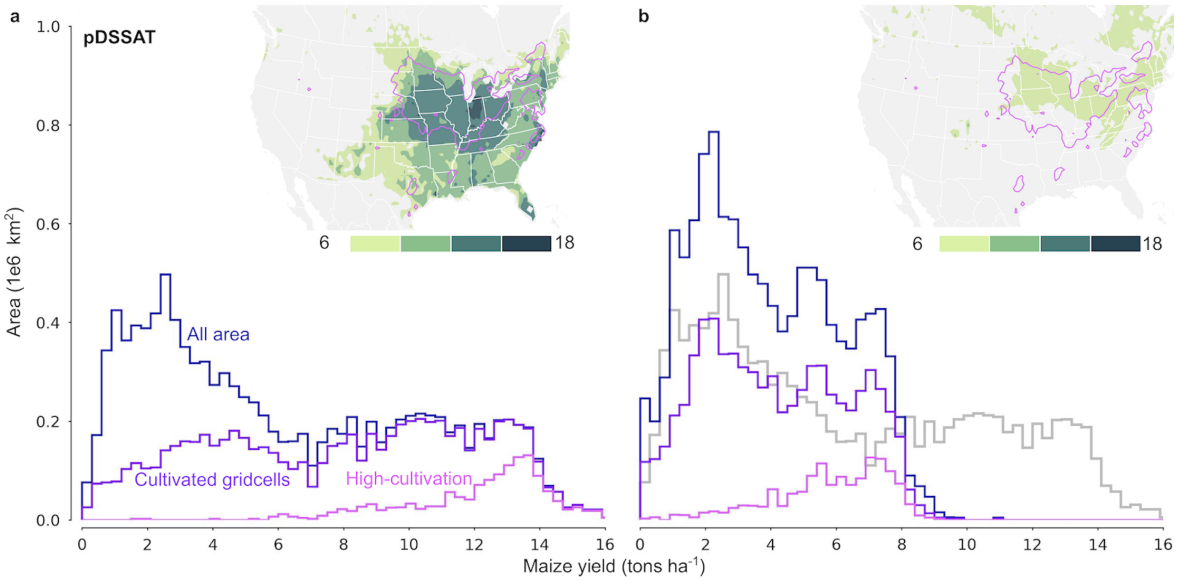


Figure S55: Same convention as Figure 4.2, except now for the pDSSAT model with high [CO_2] of 810 ppm. Simulations are the A0 growing season specification.

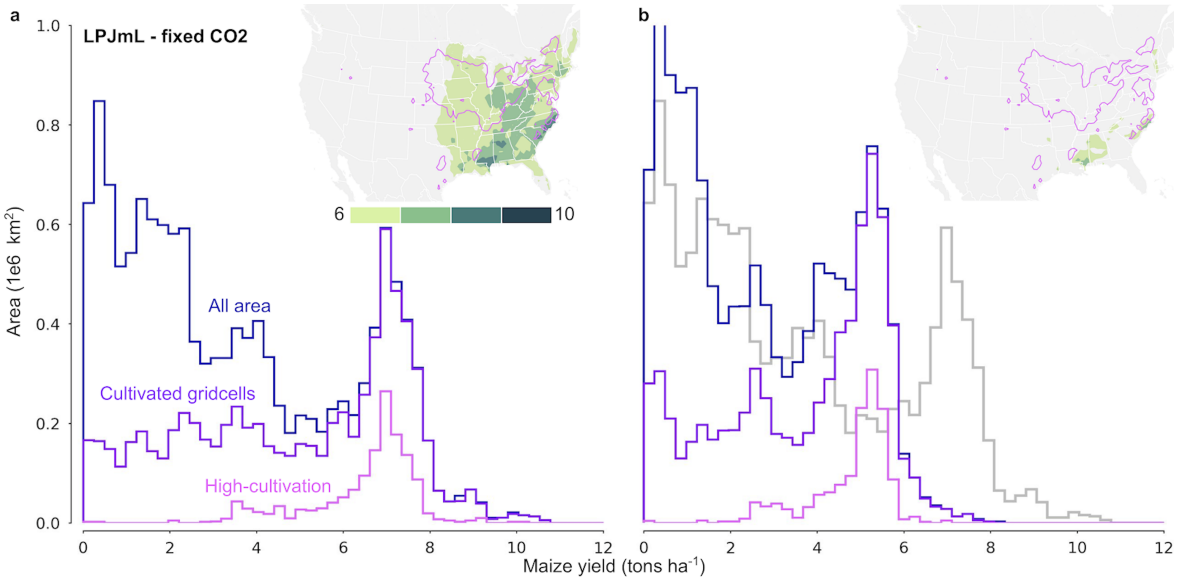


Figure S56: Same convention as Figure 4.2, except now for LPJmL (GGCMI Phase 2). Simulations are the A0 growing season specification.

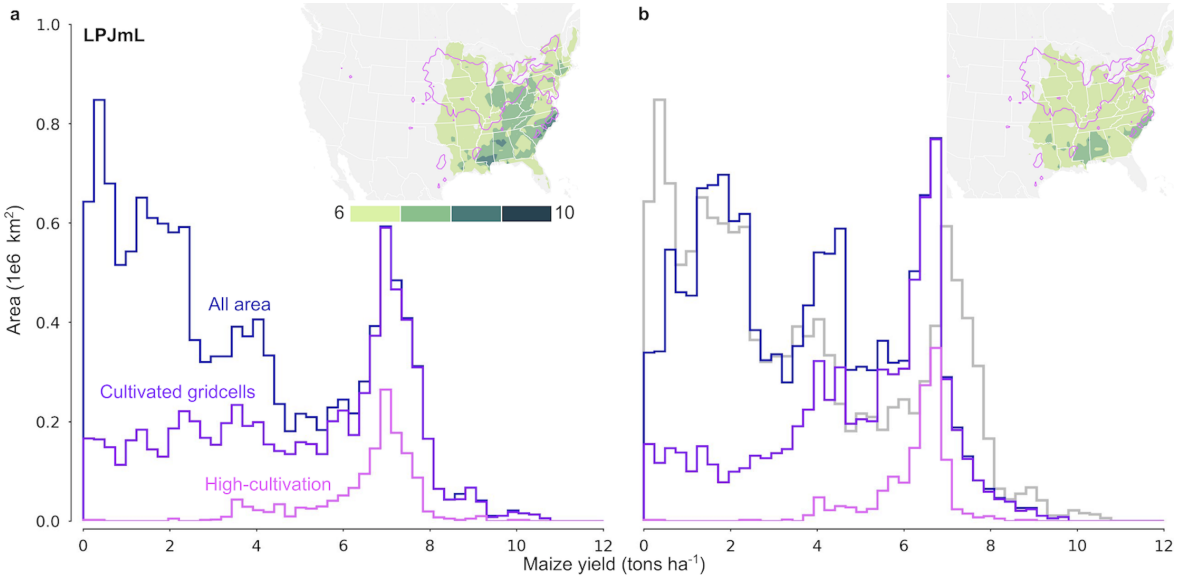


Figure S57: Same convention as Figure 4.2, except now for LPJmL (GGCMI Phase 2) with high [CO_2] of 810 ppm. Simulations are the A0 growing season specification.

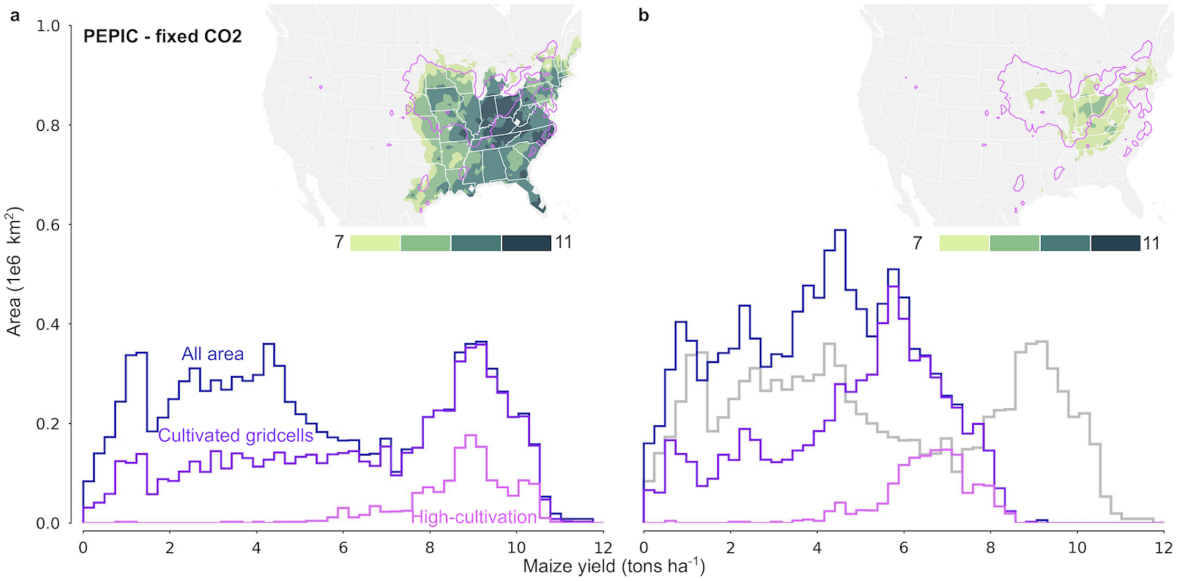


Figure S58: Same convention as Figure 2, except now for PEPIC. Simulations are the A0 growing season specification.

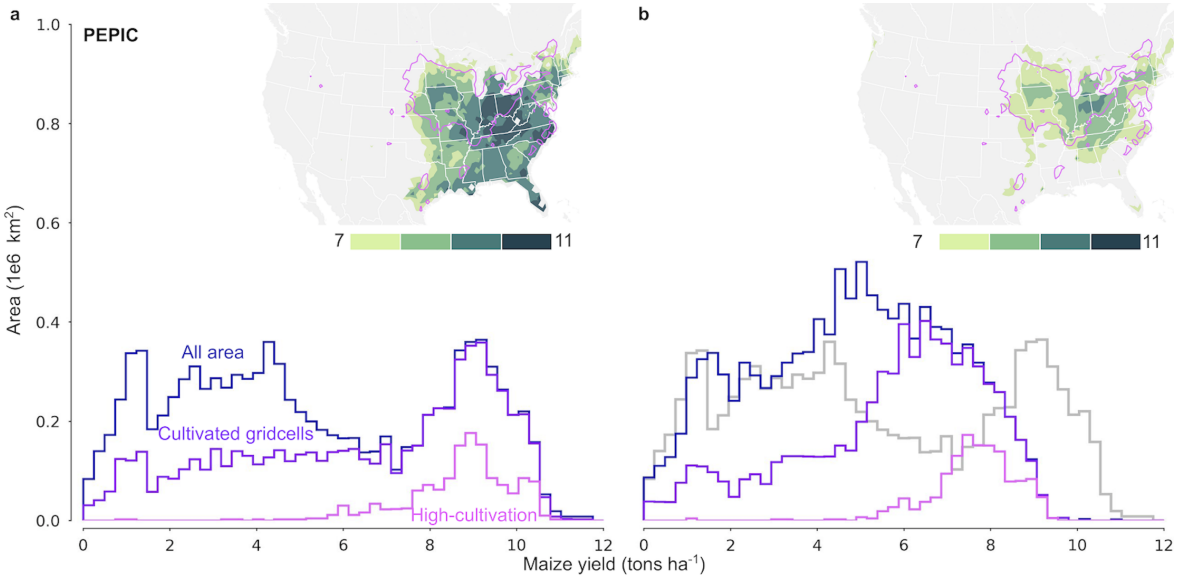


Figure S59: Same convention as Figure 4.2, except now for PEPIC with high [CO_2] of 810 ppm. Simulations are the A0 growing season specification.

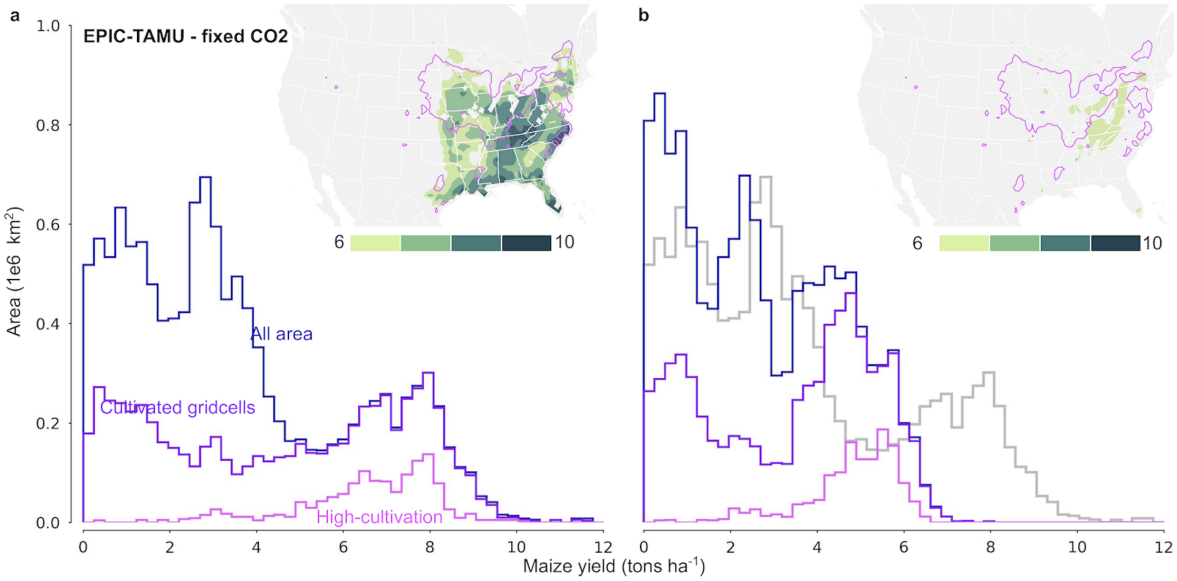


Figure S60: Same convention as Figure 4.2, except now for EPIC-TAMU. Simulations are the A0 growing season specification.

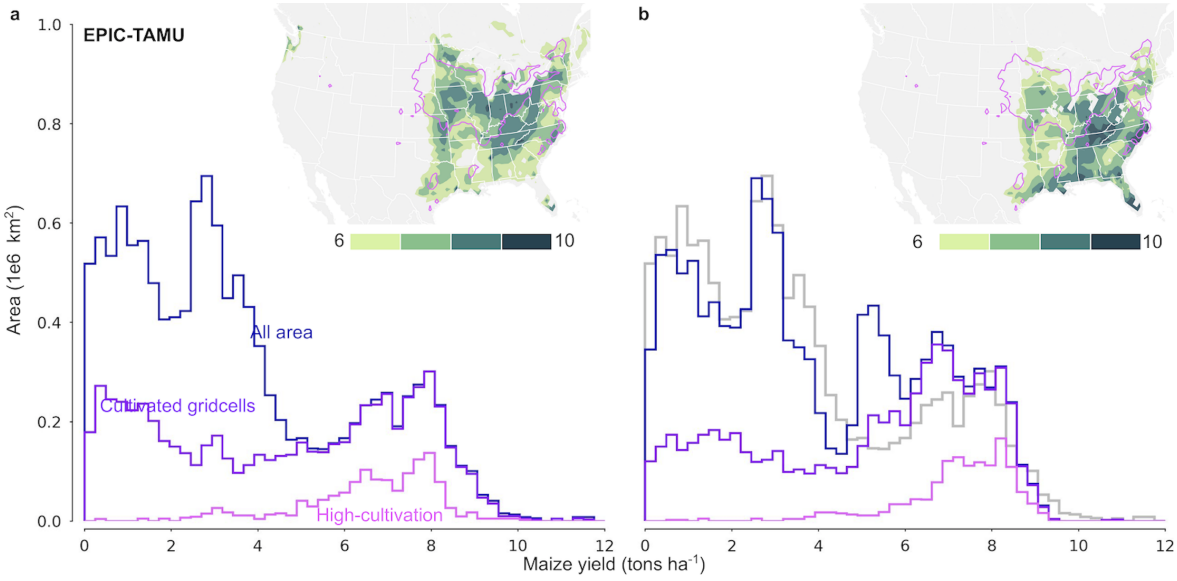


Figure S61: Same convention as Figure 4.2, except now for EPIC-TAMU with high [CO_2] of 810 ppm. Simulations are the A0 growing season specification.

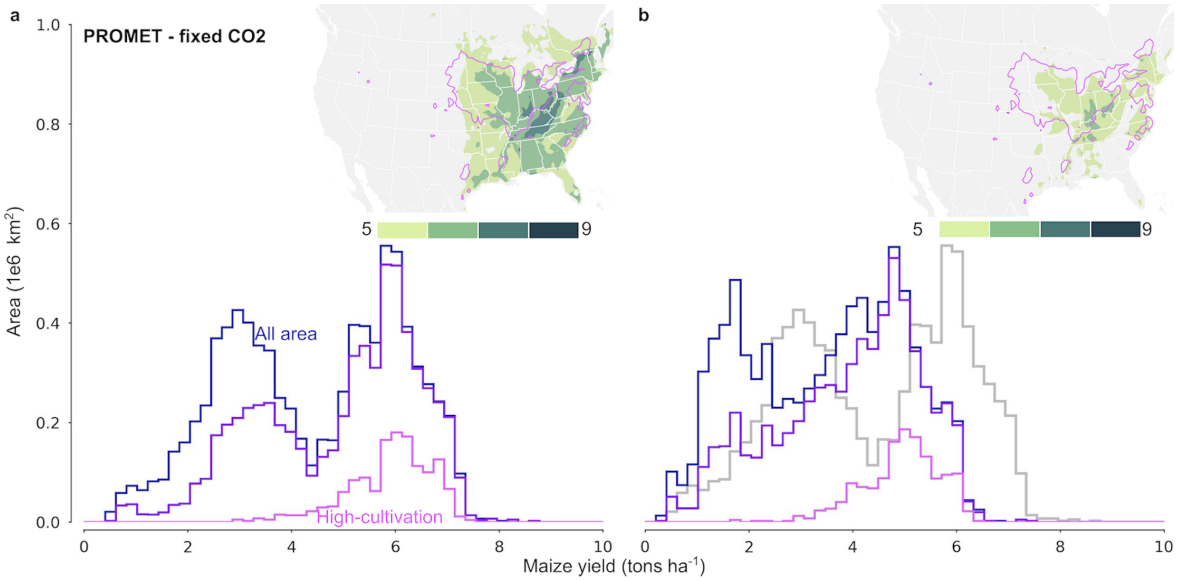


Figure S62: Same convention as Figure 4.2, except now for PROMET. Simulations are the A0 growing season specification.

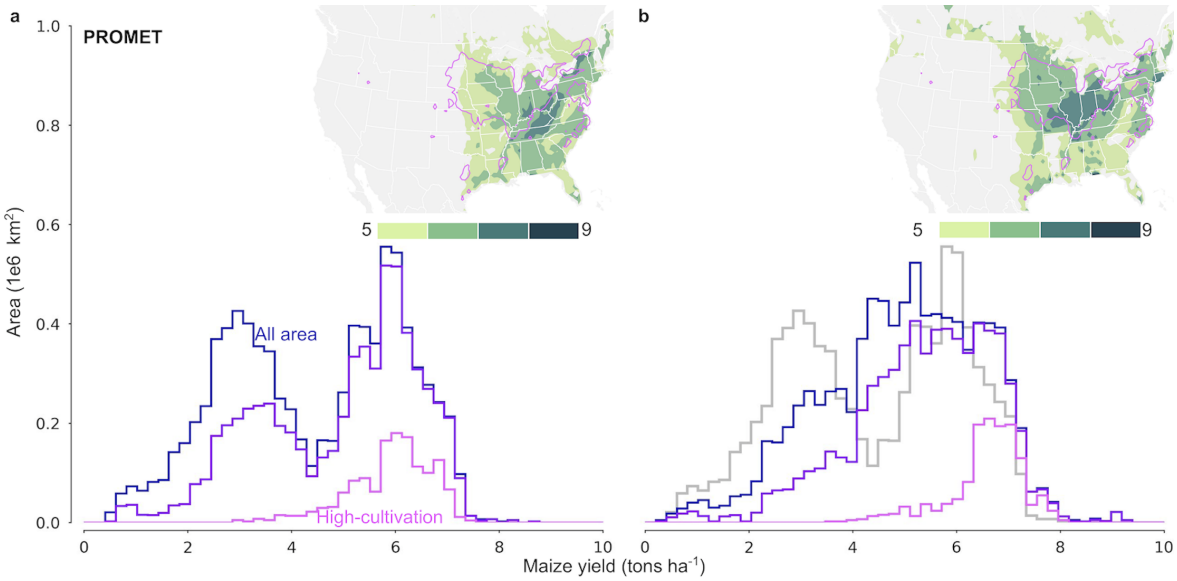


Figure S63: Same convention as Figure 4.2, except now for PROMET with high [CO_2] of 810 ppm. Simulations are the A0 growing season specification.

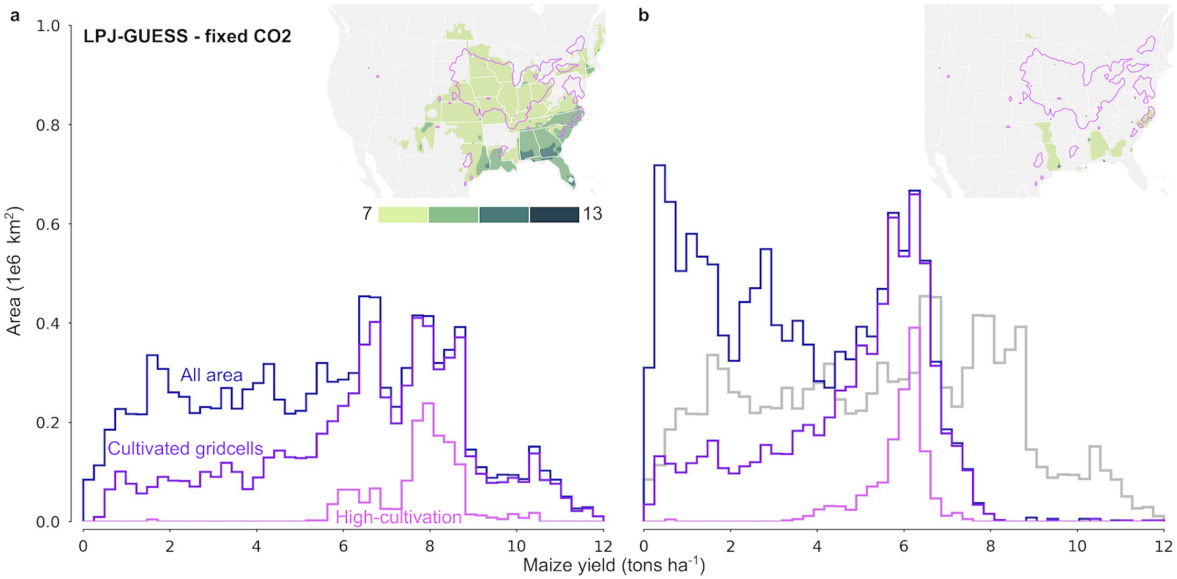


Figure S64:]
Same convention as Figure 4.2, except now for LPJ-GUESS. Simulations are the A0 growing season specification.

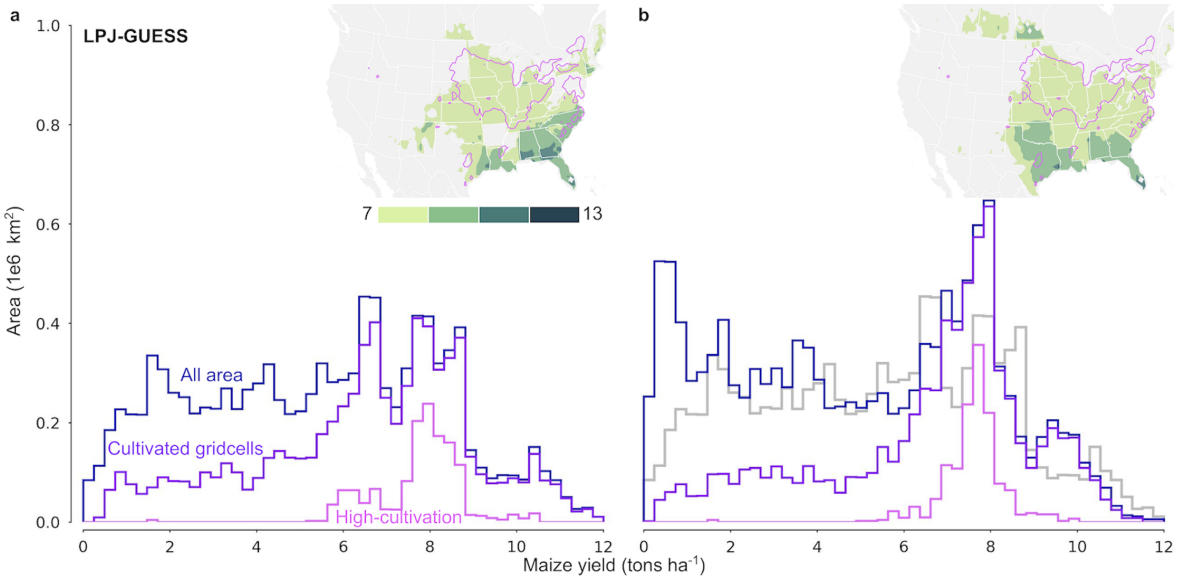


Figure S65: Same convention as Figure 4.2, except now for LPJ-GUESS with high [CO_2] of 810 ppm. Simulations are the A0 growing season specification.

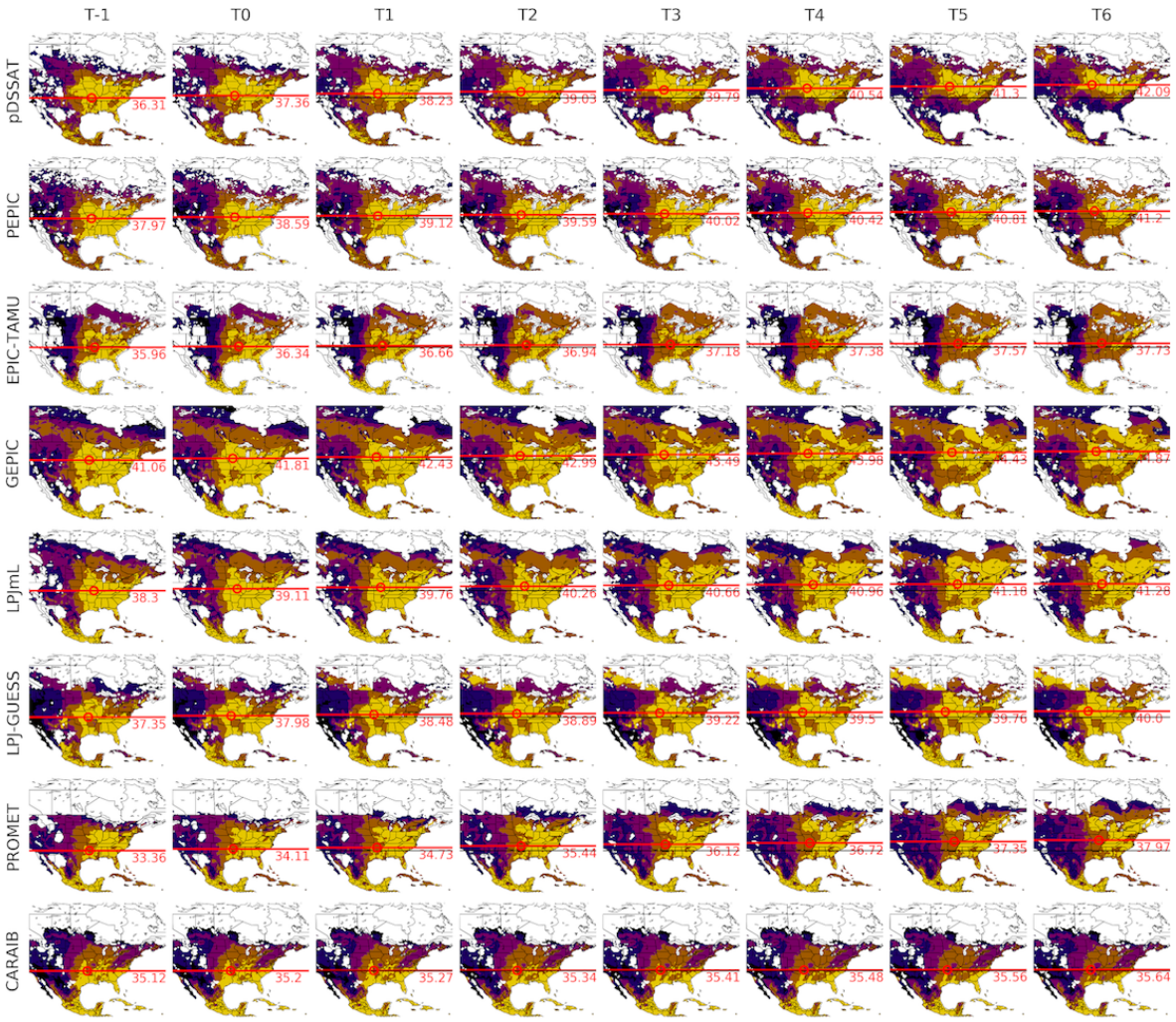


Figure S66: Yield centroid for each model for maize under A0 growing season with [CO_2] of 360 ppm under various degrees of warming (1°C, 2°C, etc). Centroid latitude noted in red. Heatmap shows yield with warmer colors indicating higher yield. Colormaps normalized across models. Simulations are the A0 growing season specification.

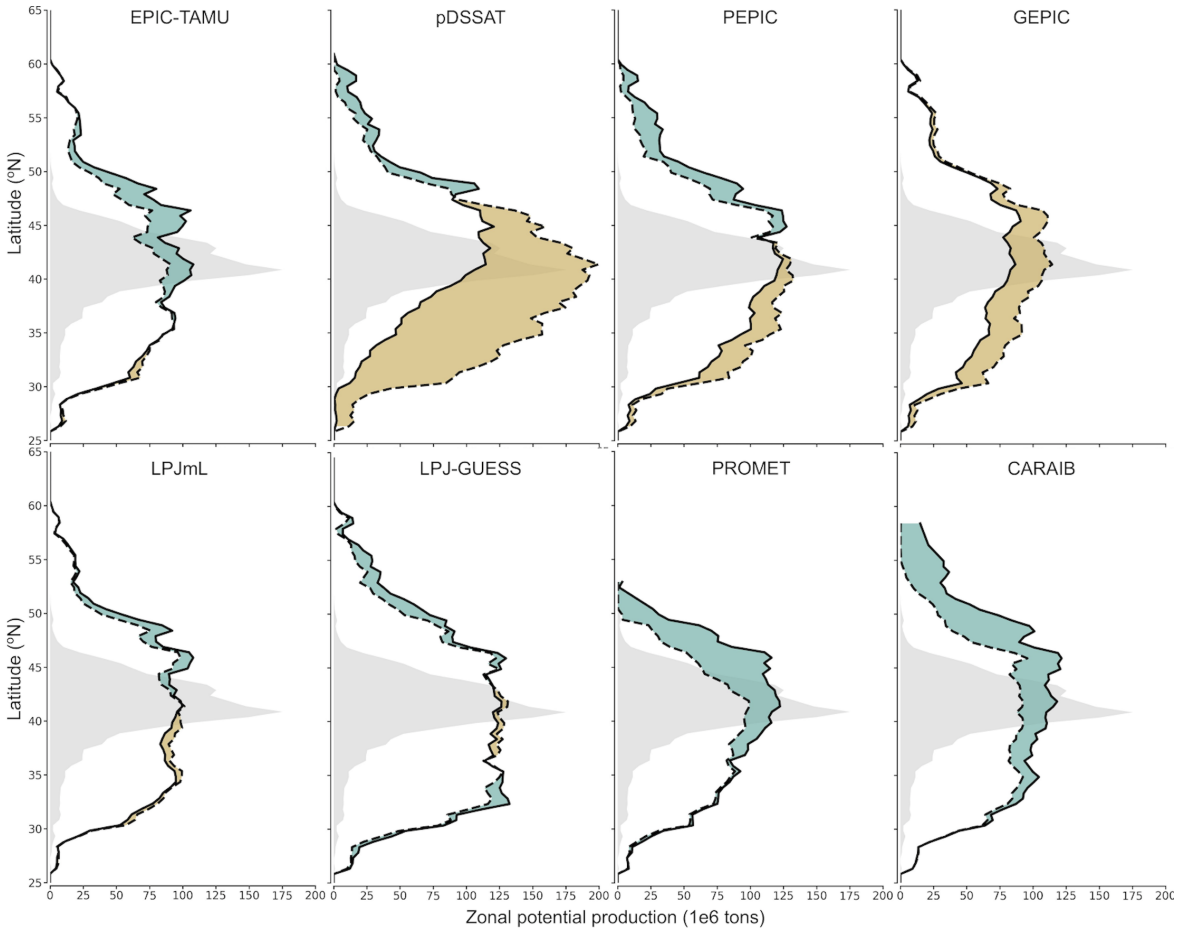


Figure S67: Same convention as above (T+6°C, [CO_2]=810ppm), now for individual models for maize in North America only. Simulations are the A0 growing season specification.

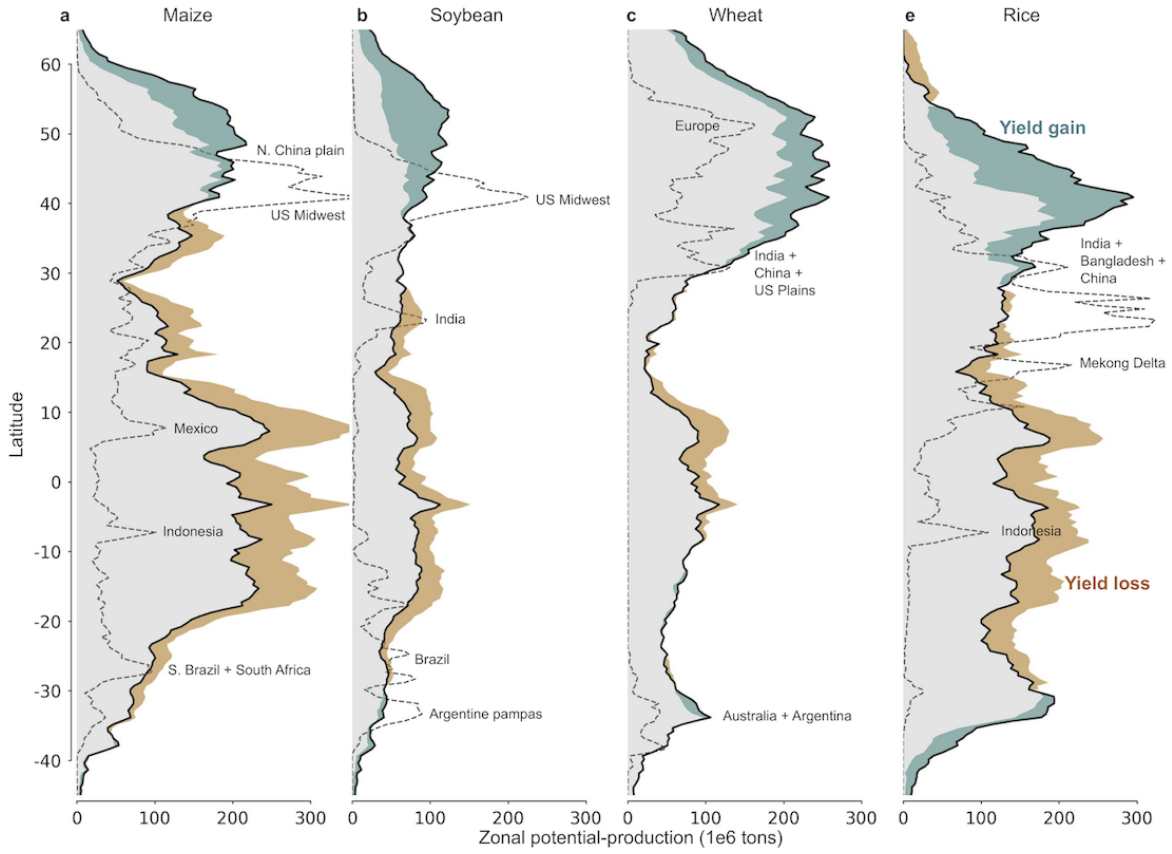


Figure S68: Same convention as Figure 4.3, now with simulations at T+6°C and [CO_2] of 810 ppm. Simulations are the A0 growing season specification.

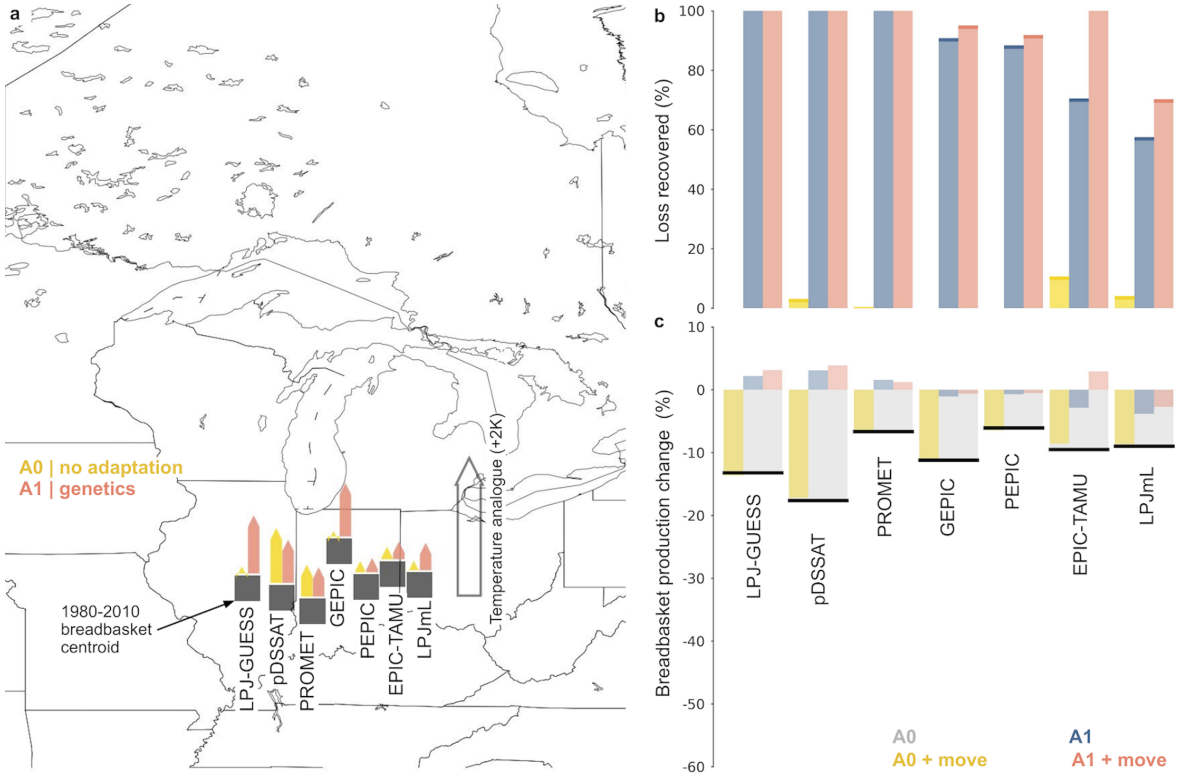


Figure S69: Same convention as Figure 4.4, now for 2°C warming. Although effects are smaller, the shifting patterns are the same.

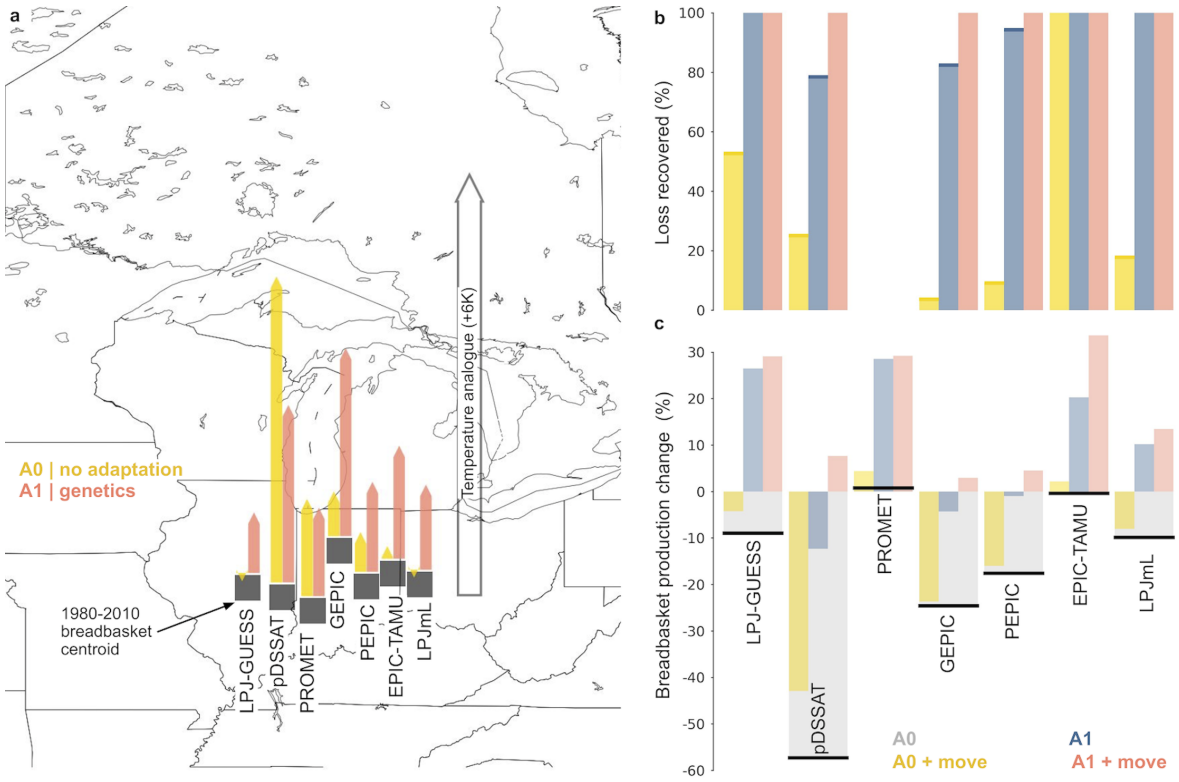


Figure S70: Same convention as Figure 4.4, now with [CO_2] at 810 ppm. Shifting patterns are the same as at low [CO_2]. PROMET experiences no losses under 6°C warming and 810 ppm.

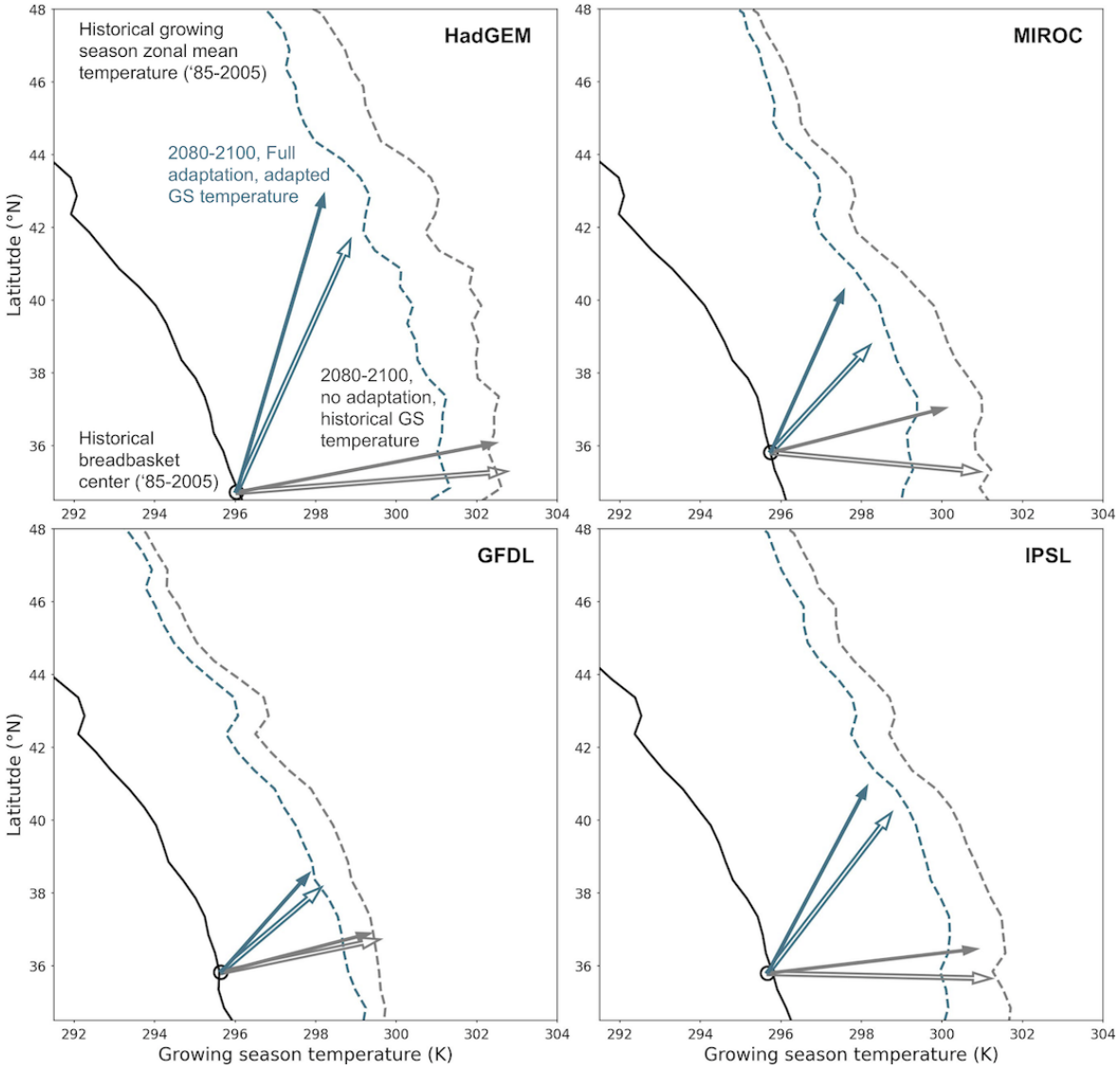


Figure S71: Climate model uncertainty. Same convention as Figure 5, now for the other climate models. Shown are two separate calculations for mean growing season temperature: one with the planting and harvest date fixed from the historical and one with the actual planting and harvest dates in the model under warming. Trajectories with dynamic [CO_2], according to the RCP, are shown as open arrows.

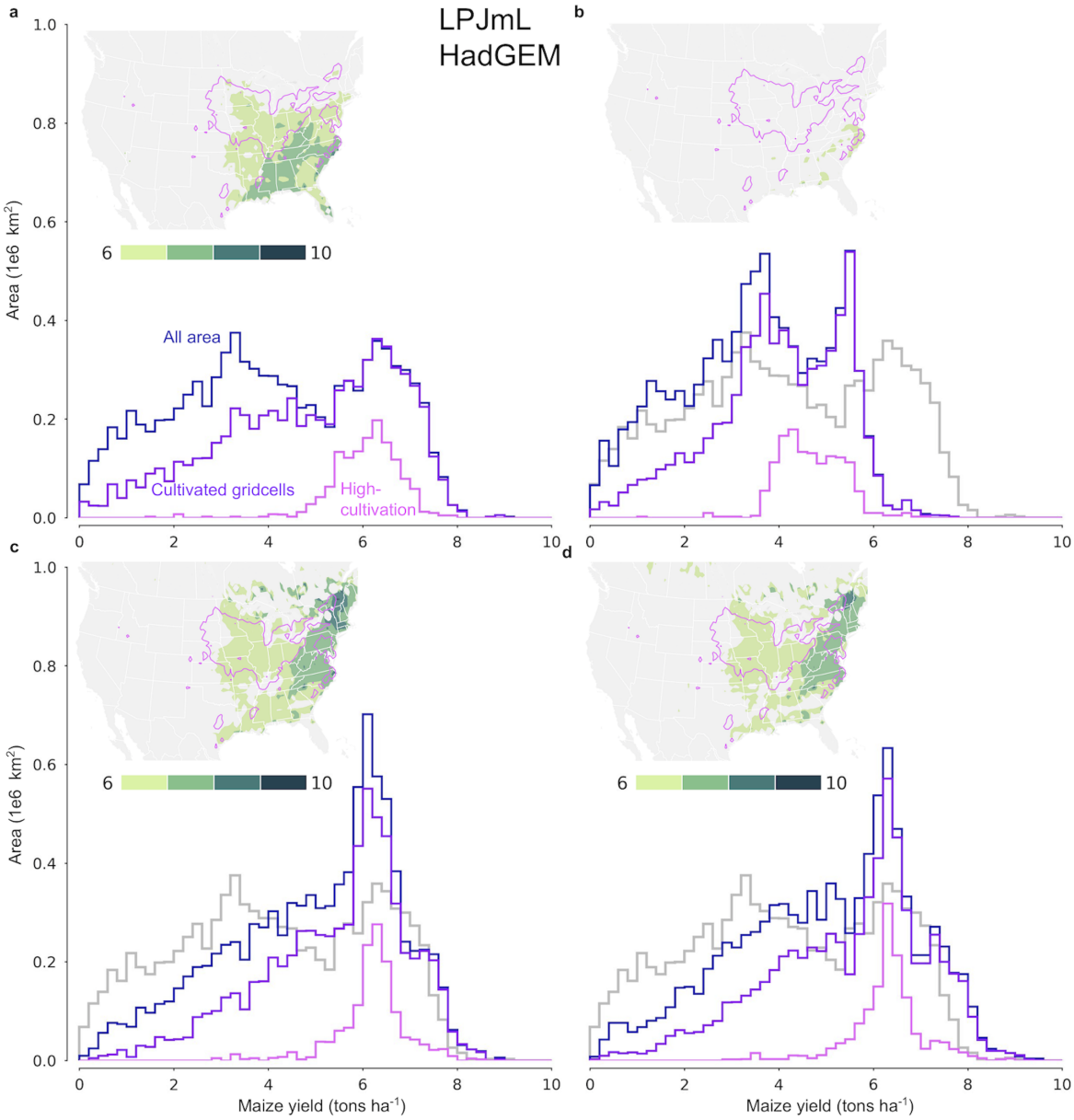


Figure S72: Same convention as Figure 4.2, except now for LPJmL transient runs driven by the HadGEM climate. (a) historical conditions (1985-2005) (b) end of the century (2080-2100) for RCP8.5 (c) end of the century with planting and harvest adaptation (d) end of century with planting and harvest adaptation and Iowa soils and radiation uniformly applied everywhere.

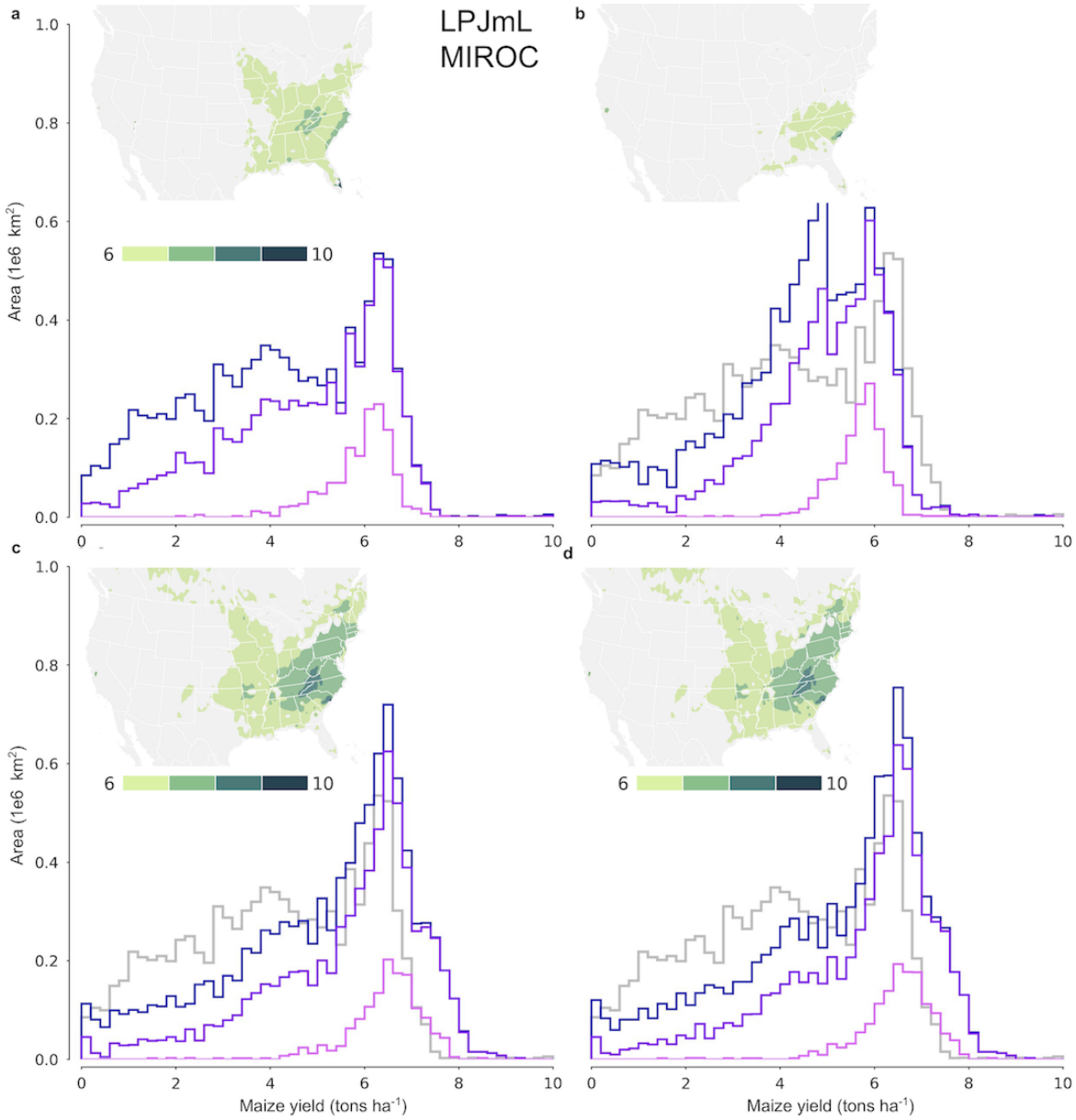


Figure S73: Same convention as above, except now for LPJmL transient runs driven by MIROC5.

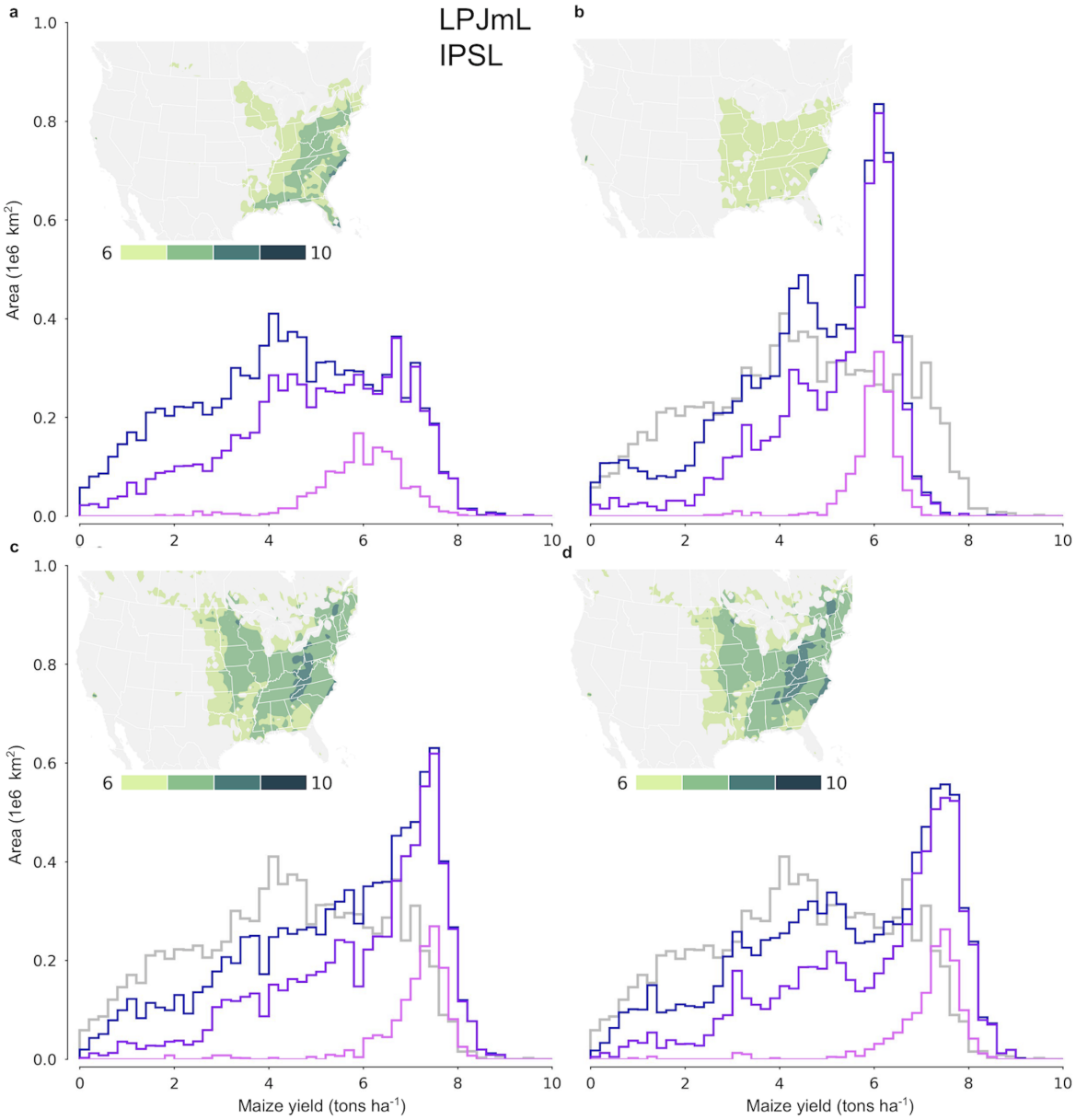


Figure S74: Same convention as above, except now for LPJmL transient runs driven by IPSL.

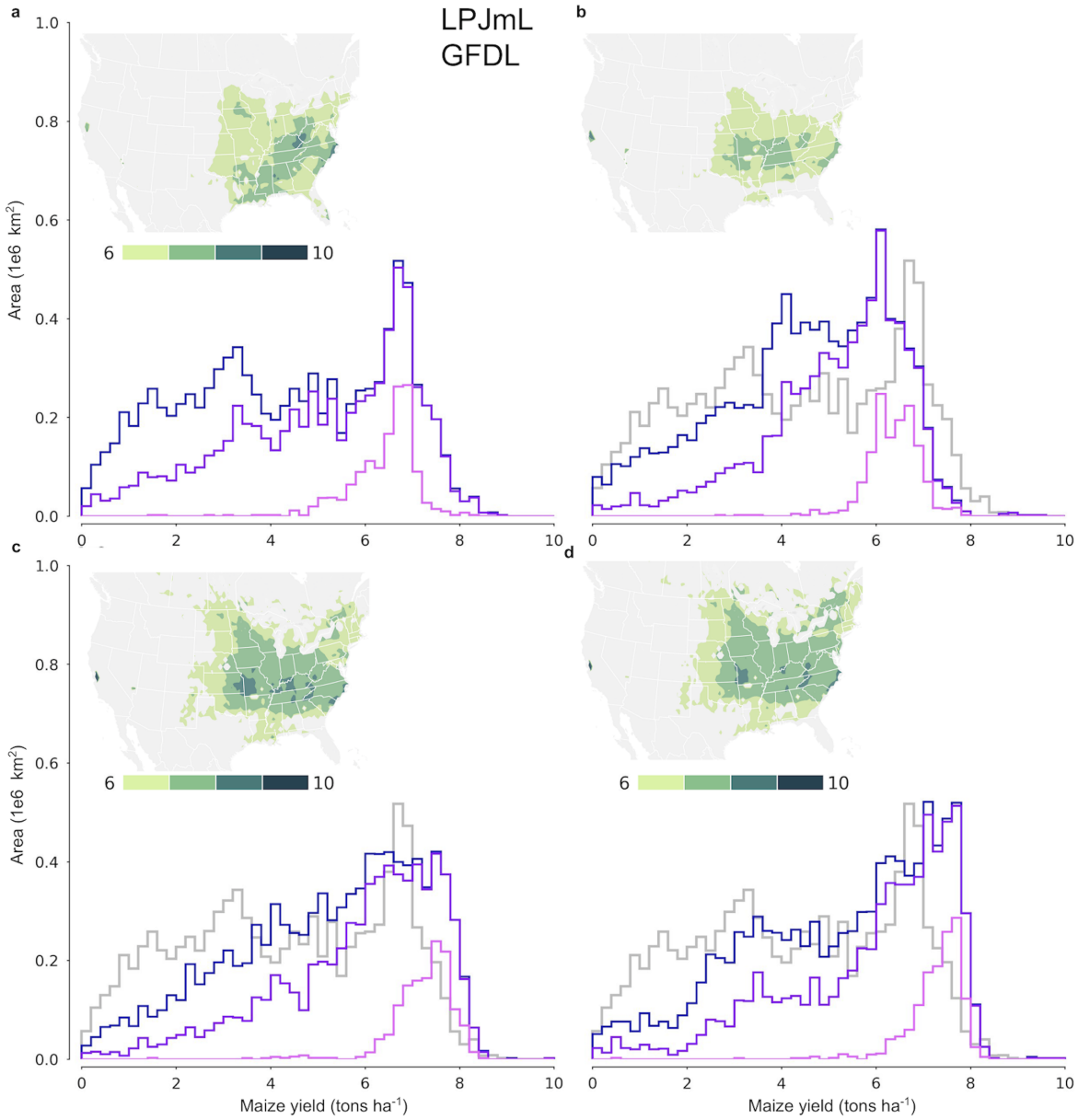


Figure S75: Same convention as above, except now for LPJmL transient runs driven by GFDL.

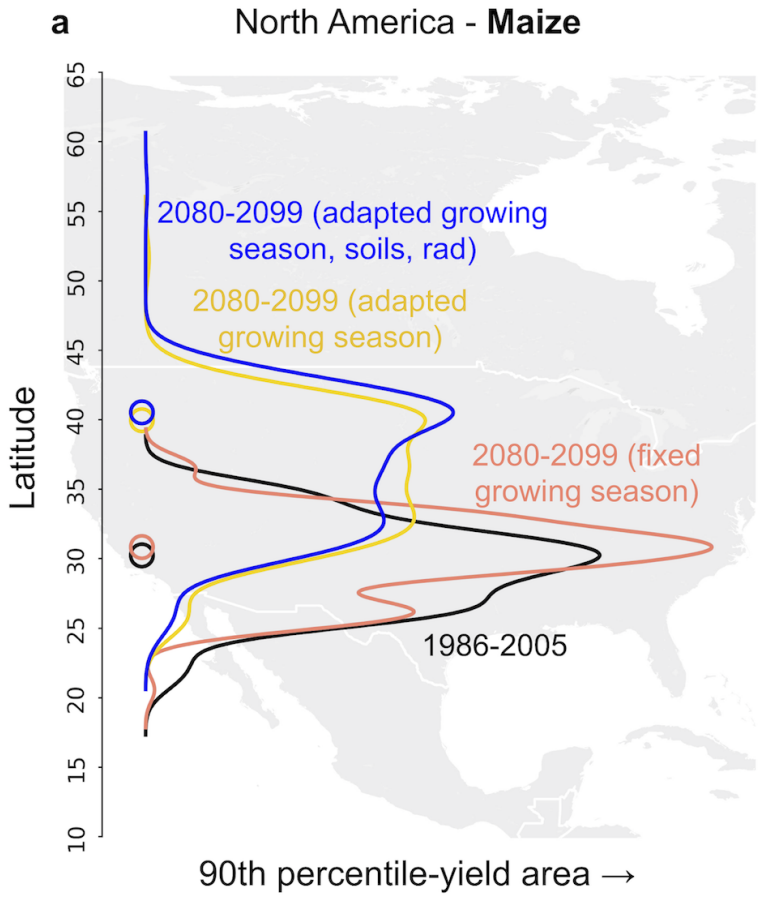


Figure S76: Latitudinal distribution of the highest yielding grid cells in the LPJmL simulation under different growing season adaptation conditions driven by the HADGEM-2 model RCP8.5. Peak latitude shown by the open circle. Growing season adaptations are critical to shift the highest-yielding land. Soils and radiation play a lesser role.

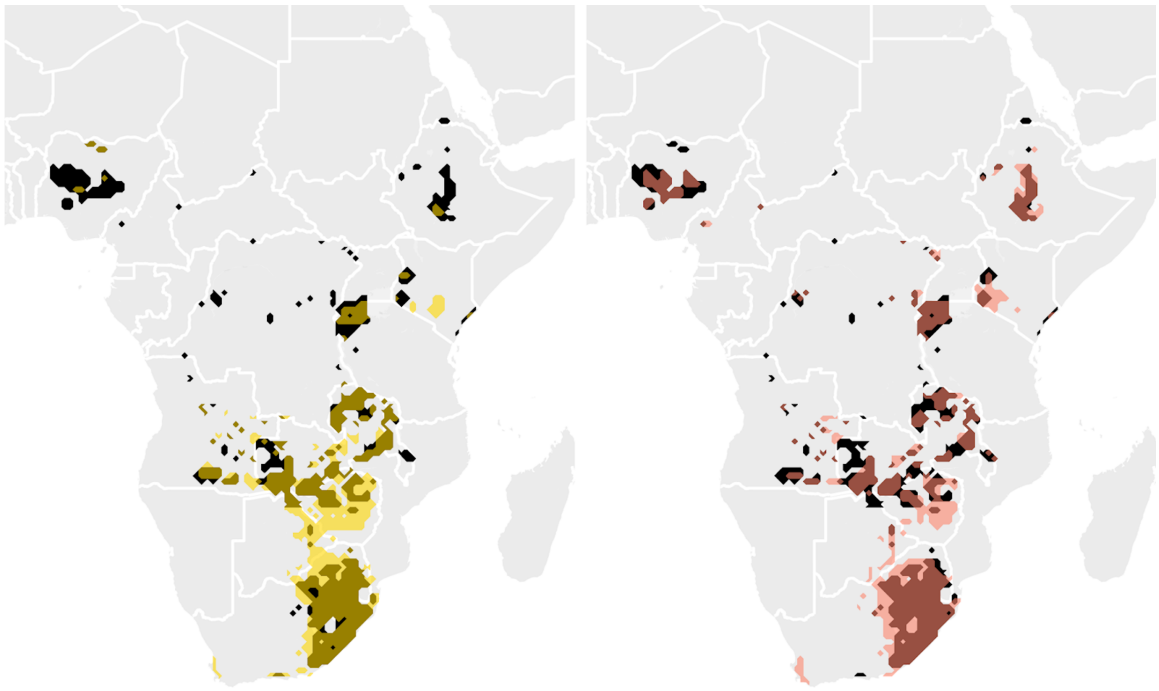


Figure S77: Repeat of Figure 4.6 for maize in Africa. Gold = no adaptation. Salmon = full adaptation. The main maize breadbasket in South Africa has nowhere to shift to under warming. Yields in Nigeria and Ethiopia are more resilient with growing season adaptation measures. The simulated maize breadbasket in Africa in less contiguous than other breadbaskets.

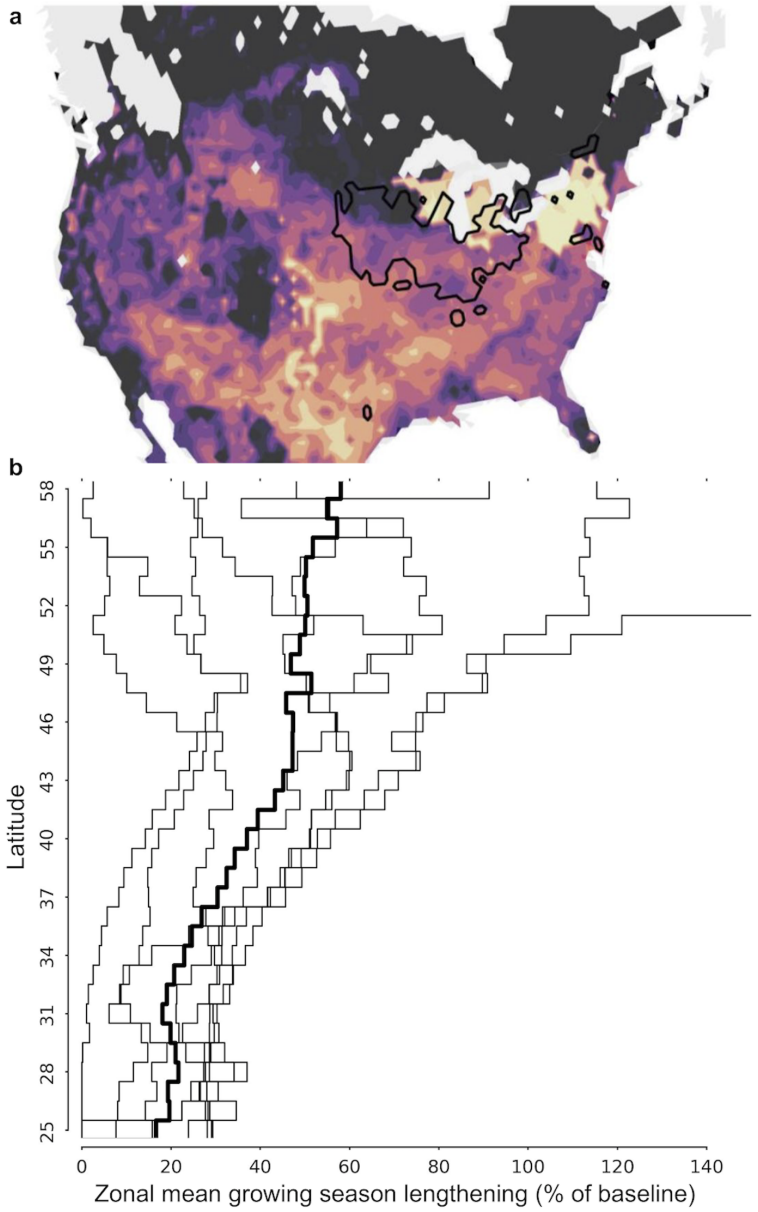


Figure S78: Latitudinal uncertainty in growing seasons changes with warming for maize in North America. Lines indicate the difference in growing season length between A1 and A0 under T+6°C of warming as a percentage of baseline. Zonal mean in North America across agricultural land. Thin gray lines indicated individual models and the thick black line is the ensemble mean. Good model agreement in the mid latitudes where the current breadbasket is centered disappears north of 46 and especially north of 49 latitude.

S4 Supplemental Materials for Chapter 5

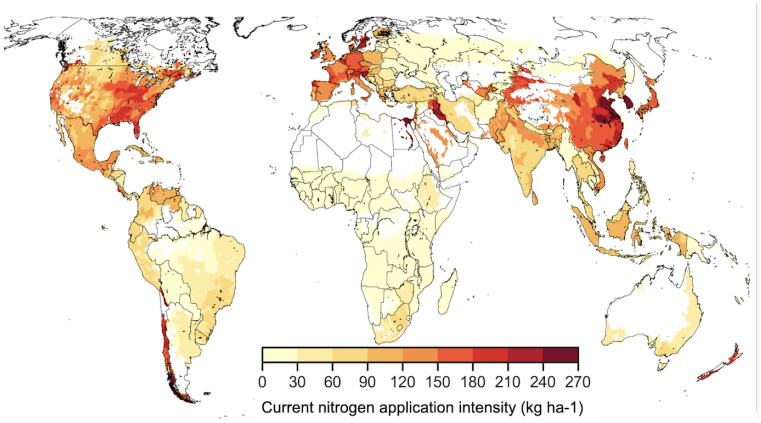


Figure S79: Current global nitrogen applications

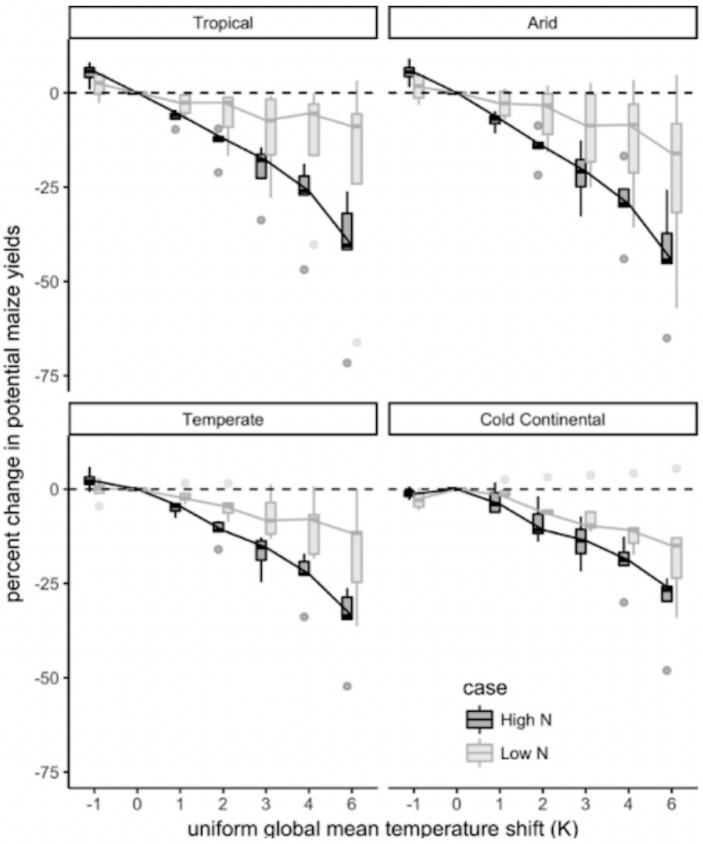


Figure S80: Nitrogen-driven climate impact miss-estimation by main KG region.