

Supplemental Material
**The GGCMI phase II emulators: global gridded crop model
responses to changes in CO₂, temperature, water, and nitrogen
(version 1.0)**

James Franke^{1,2}, Christoph Müller³, Joshua Elliott^{2,4}, Alexander Ruane⁵, Jonas Jägermeyr^{3,2,4,5},
Abigail Snyder⁶, Marie Dury⁷, Pete Falloon⁸, Christian Folberth⁹, Louis François⁷, Tobias Hank¹⁰,
Cesar Izaurrealde^{11,12}, Ingrid Jacquemin⁷, Curtis Jones¹¹, Michelle Li^{2,13}, Wenfeng Liu^{14,15},
Stefan Olin¹⁶, Meridel Phillips^{5,17}, Thomas Pugh^{18,19}, Ashwan Reddy¹¹, Karina Williams⁸,
Ziwei Wang^{1,2}, Florian Zabel¹⁰, 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. 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, Institut d’Astrophysique et de Géophysique, 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, F-91191 Gif-sur-Yvette, 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.

Contents

1 Experiment simulation sampling in variable space	2
2 Cultivation areas	3
3 Yield response for A1 simulations	4
4 Normalized error for other cases	6
5 Comparison to transient climate run simulation	8
6 Emulator products	9
7 Reduced specification (23-term) emulator examples	12
8 Yield Responses for other crops and models	17
9 Normalized Error for all models	19
10 Cross validation error for all models	23

1 Experiment simulation sampling in variable space

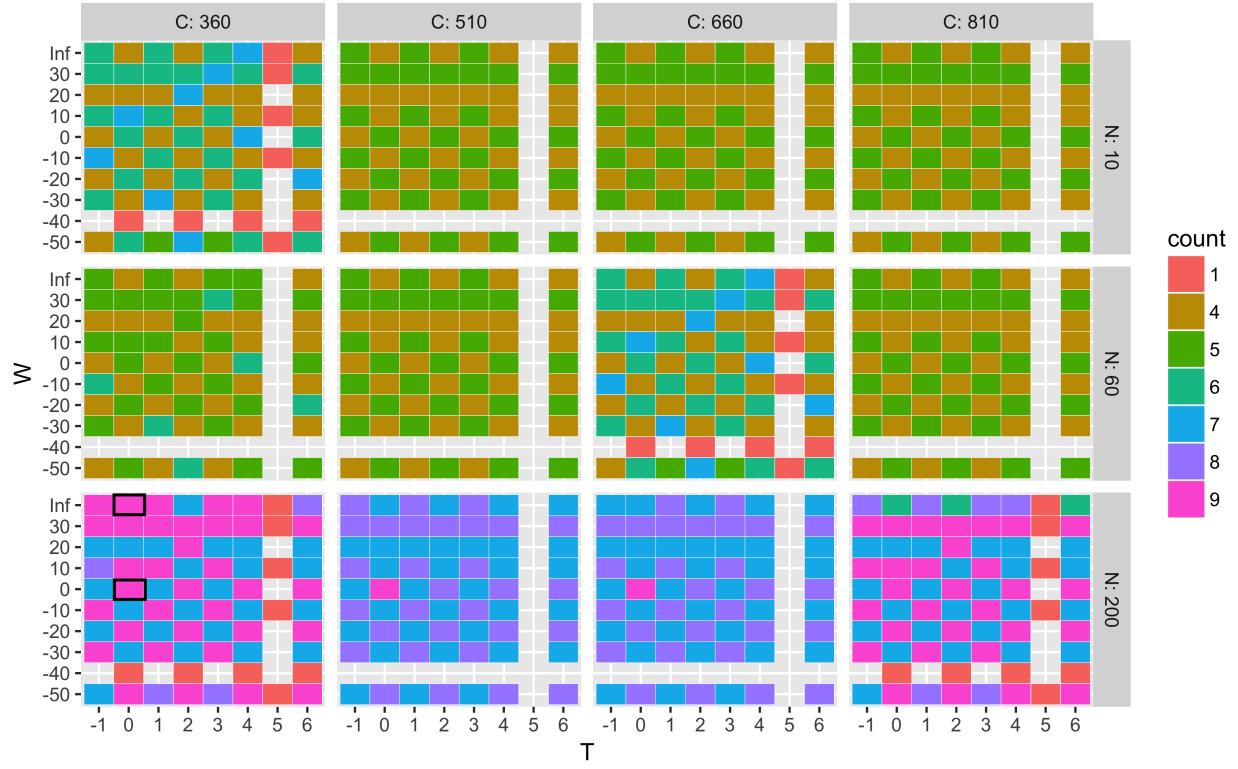


Figure S1: Heatmap illustrates number of model simulations provided for each of the scenarios in the variable space. The max number is 9, the number of models included in the emulator analysis (excluding three models not included in the emulator analysis). Normalized error calculations are run over scenarios with max number of models.

2 Cultivation areas

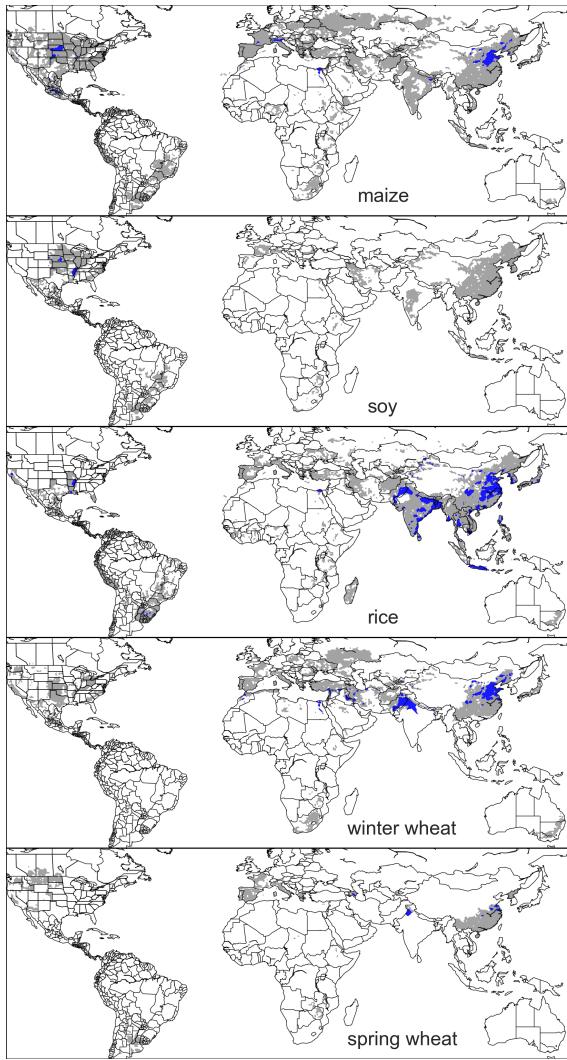


Figure S2: Presently cultivated area for irrigated crops in the real world. The blue contour area indicates grid-cells with more than 20,000 hectares of crop cultivated. The gray contour shows area with more than 10 hectares cultivated. Data from the MIRCA2000 data set for maize, rice, and soy. Winter and spring wheat areas are adapted from MIRCA2000 data and sorted by growing season

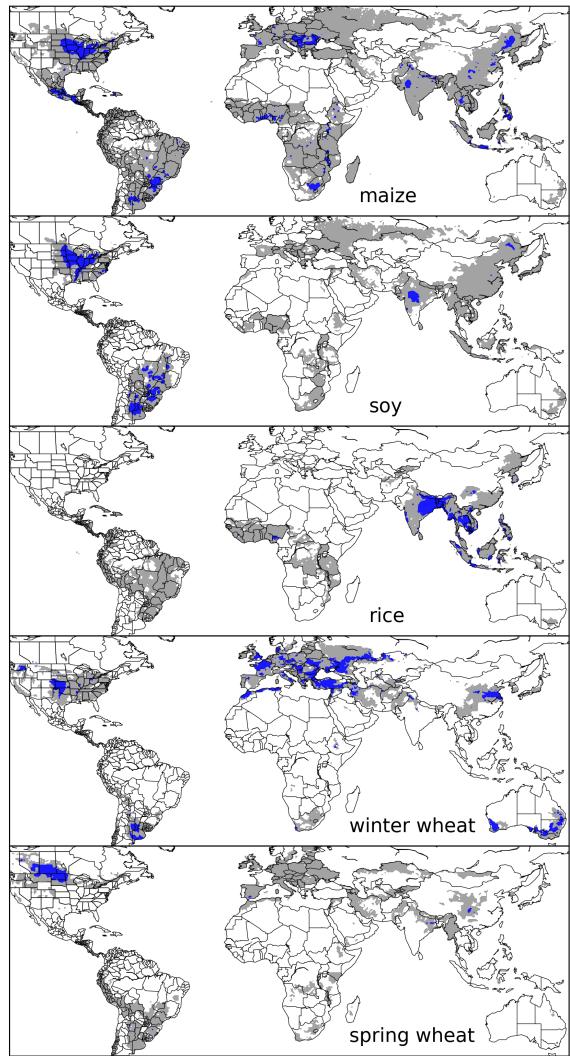


Figure S3: Presently cultivated area for rain fed crops in the real world. Conventions as in Figure S1.

3 Yield response for A1 simulations

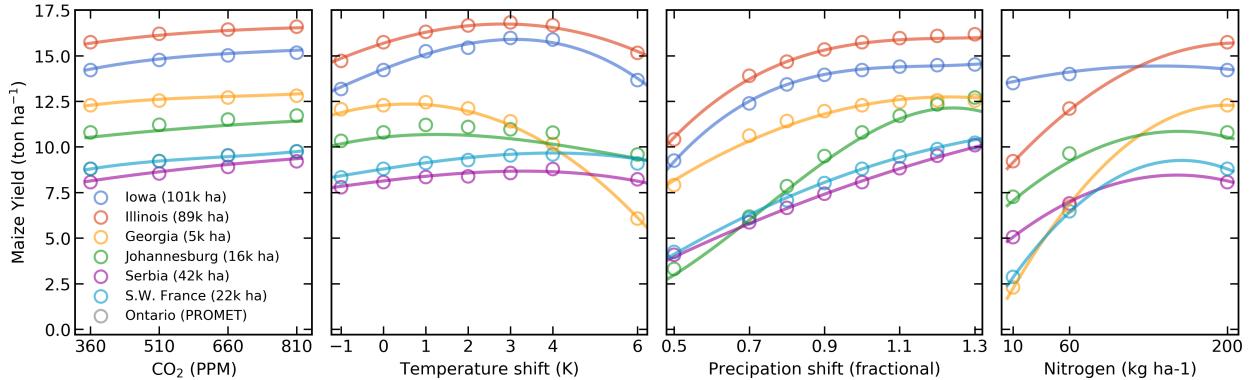


Figure S4: Illustration of spatial variations in yield response, which are successfully captured by the emulator for the A1 simulations. Panels show simulations (points) and emulations (lines) of rainfed maize in the pDSSAT model in six example locations selected to represent high-cultivation areas around 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.

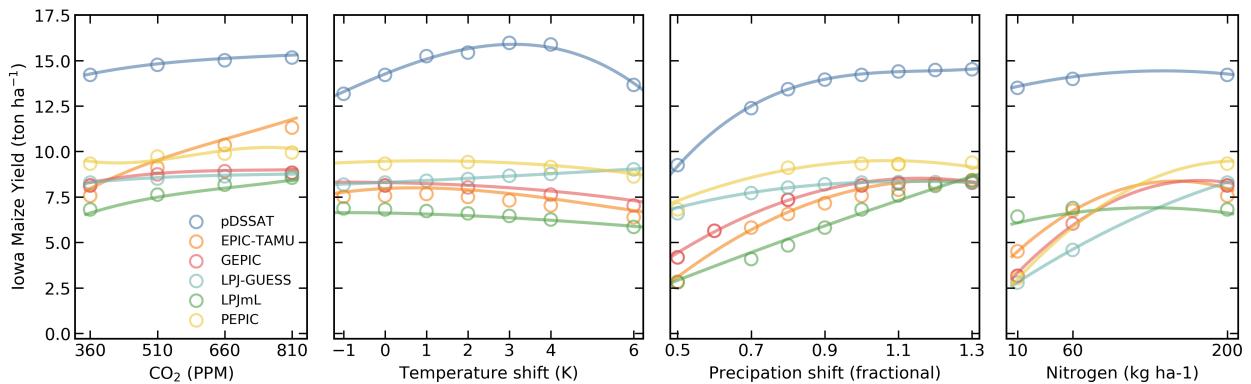


Figure S5: Illustration of variations in yield response across models, again successfully captured by the emulator. Panels show simulations and emulations from six representative GGCMI models for rainfed maize in the same Iowa grid cell shown above, with the same plot conventions. Three models (PROMET, JULES, and CARAIB) that do not simulate the nitrogen dimension are omitted for clarity. Models are uncalibrated, producing spread in absolute yields.

4 Normalized error for other cases

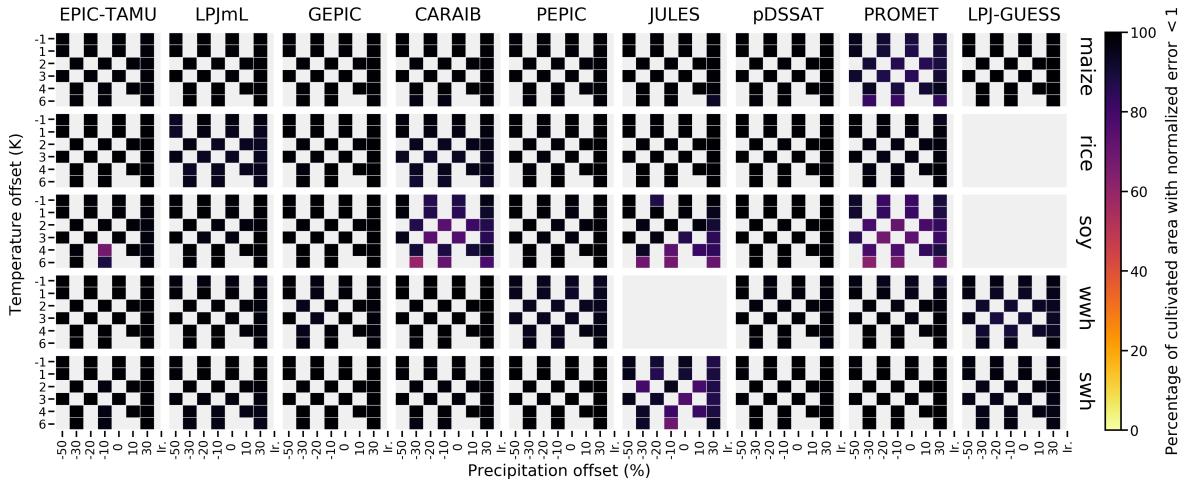


Figure S6: The fraction of currently cultivated hectares with normalized emulation error less than 1 for the $\text{CO}_2=810 \text{ ppm}$ and $200 \text{ kg N ha}^{-1} \text{ yr}^{-1}$ case for the temperature and precipitation perturbations scenarios provided by all 9 models included in the emulator analysis. The yield response is generally easy to emulate over currently cultivated areas (black regions).

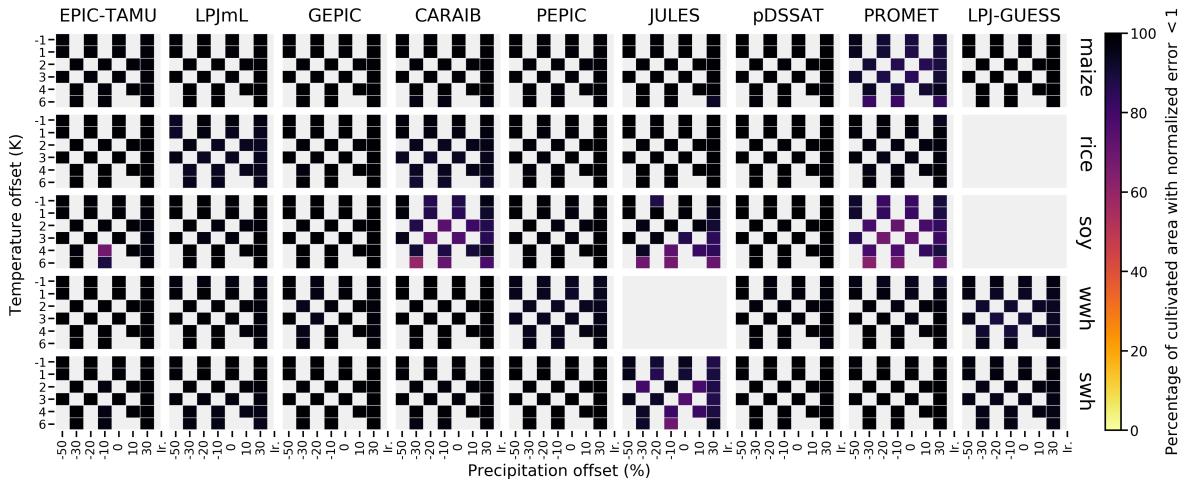


Figure S7: The fraction of currently cultivated hectares with normalized emulation error less than 1 for A1 (growing season adaptation) simulations for $\text{CO}_2=360 \text{ ppm}$ and $200 \text{ kg N ha}^{-1} \text{ yr}^{-1}$ case for the temperature and precipitation perturbations scenarios provided by all 9 models included in the emulator analysis. The yield response is generally easy to emulate over currently cultivated areas (black regions).

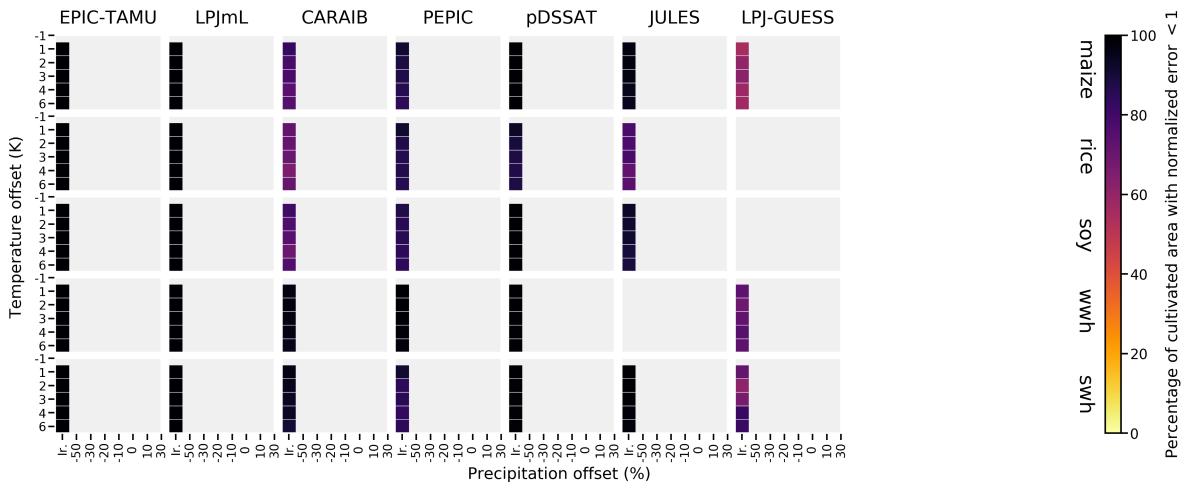


Figure S8: The fraction of currently cultivated hectares with normalized emulation error less than 1 for irrigated water demand for $\text{CO}_2=310 \text{ ppm}$ and $200 \text{ kg N ha}^{-1} \text{ yr}^{-1}$ case for the temperature and precipitation perturbations scenarios provided by all models included for irrigation water demand in the emulator analysis.

5 Comparison to transient climate run simulation

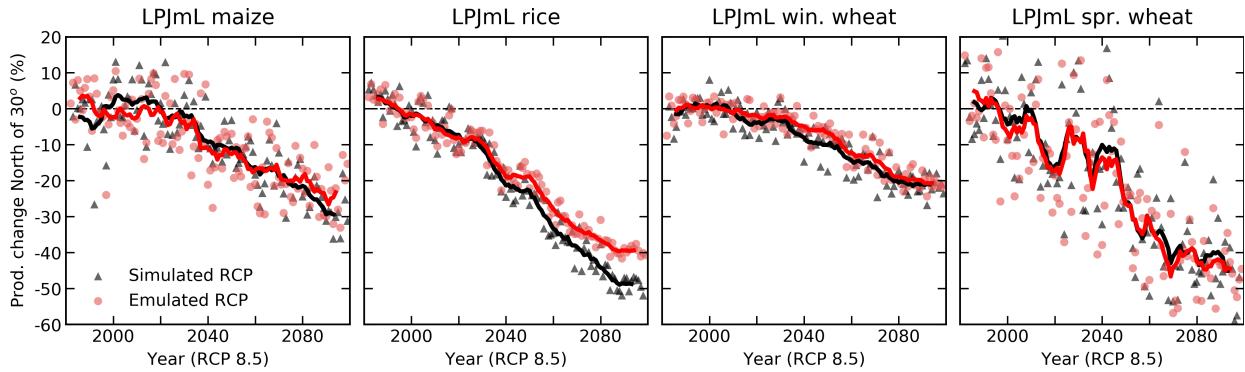


Figure S9: Same as main text Figure 9 except now only for crops north of 30N latitude.

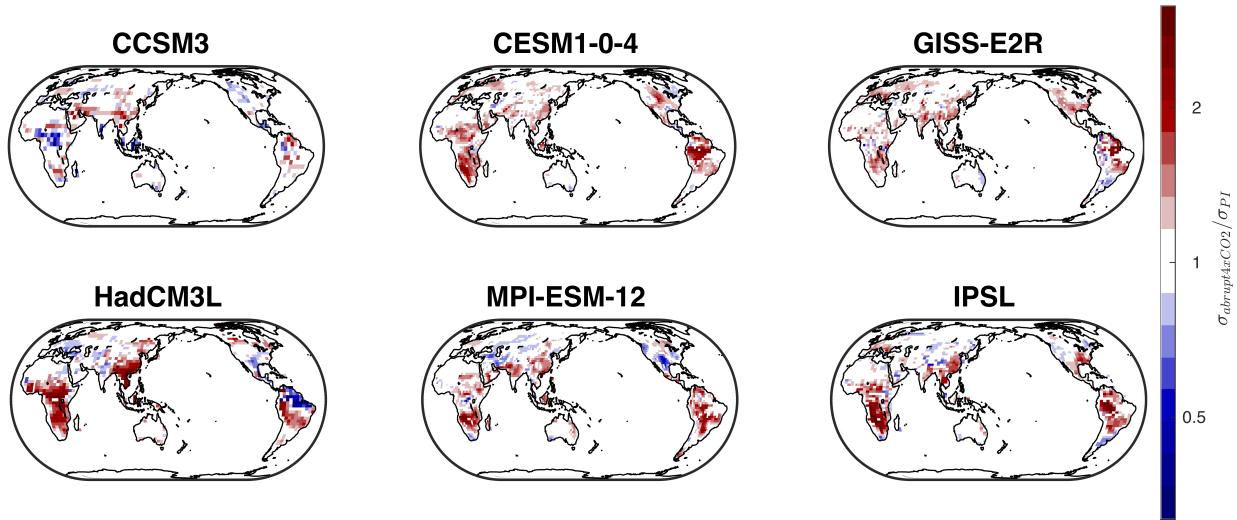


Figure S10: Change in hemispheric-summer variability for some CMIP-5 models for an abrupt 4x CO₂ forcing. Ratio in standard deviation in temperature compared to preindustrial values. The HadCM3 model shows relatively strong variability changes. Model disagreement about changes in variability is large.

6 Emulator products

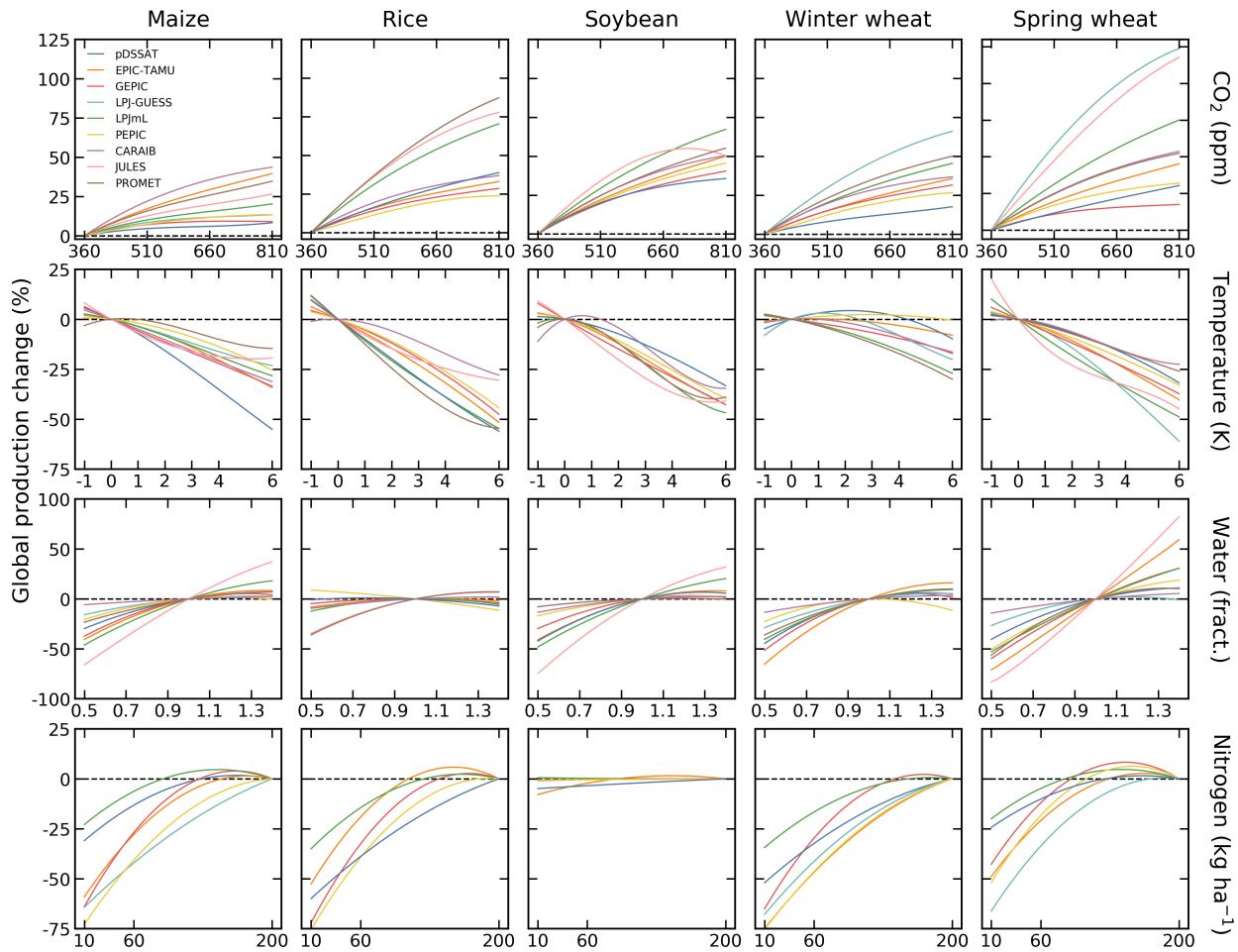


Figure S11: Same as main text Figure 10 except now each model is shown in color.

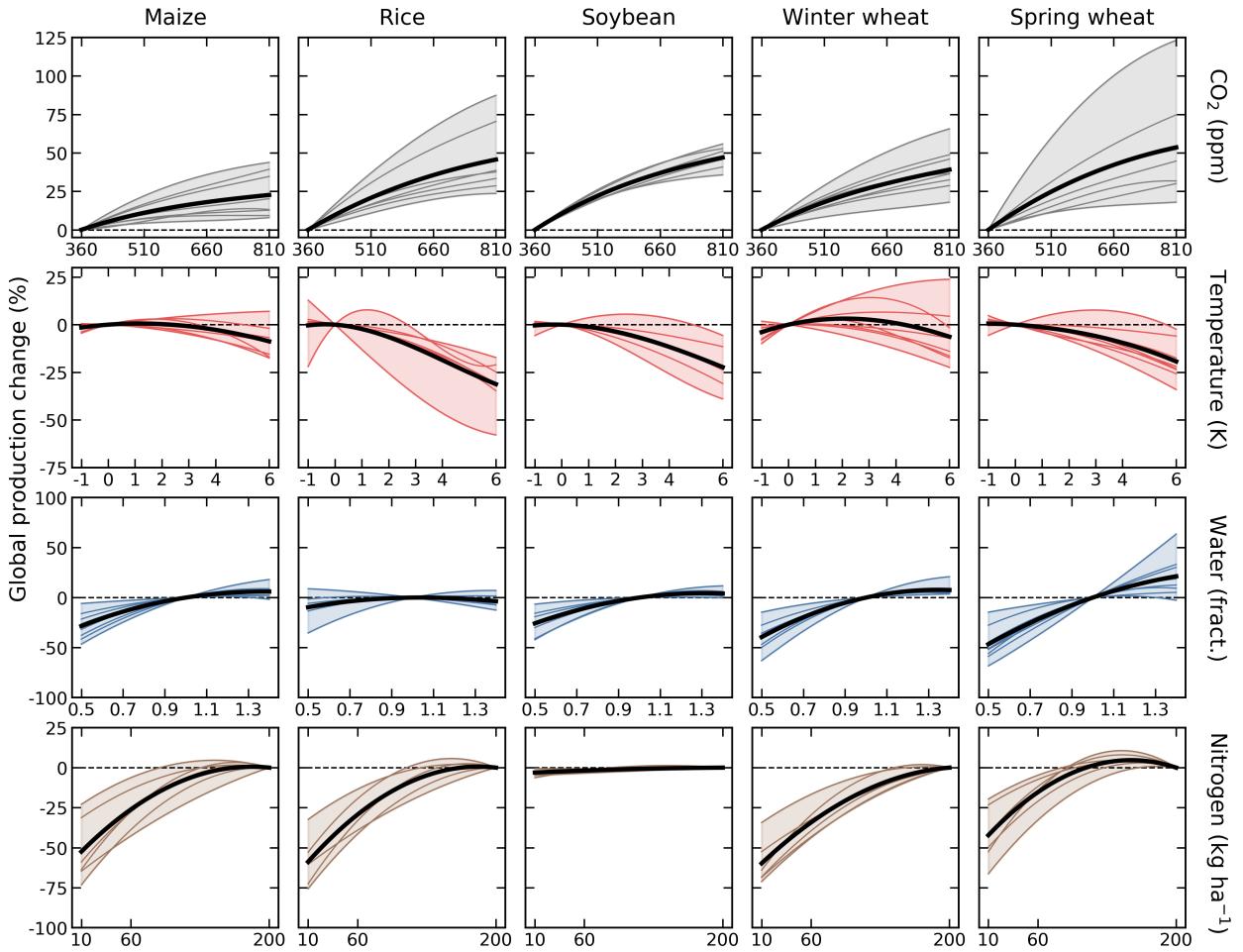


Figure S12: Emulated damage functions for rainfed crops for A1 simulations. Note: JULES does not provide A1 simulations.

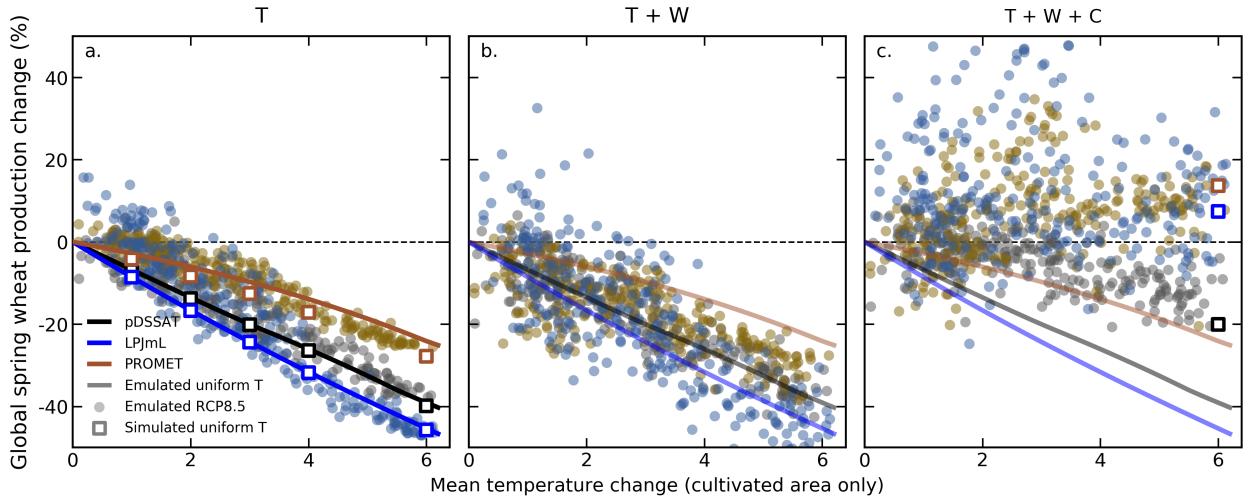


Figure S13: Illustration of the factors affecting yields in more realistic climate scenarios for spring wheat. Same convention main text Figure 11. Emulator errors in PROMET are linked to Southern China, where error are very high due to discontinuities in yield response in that region.

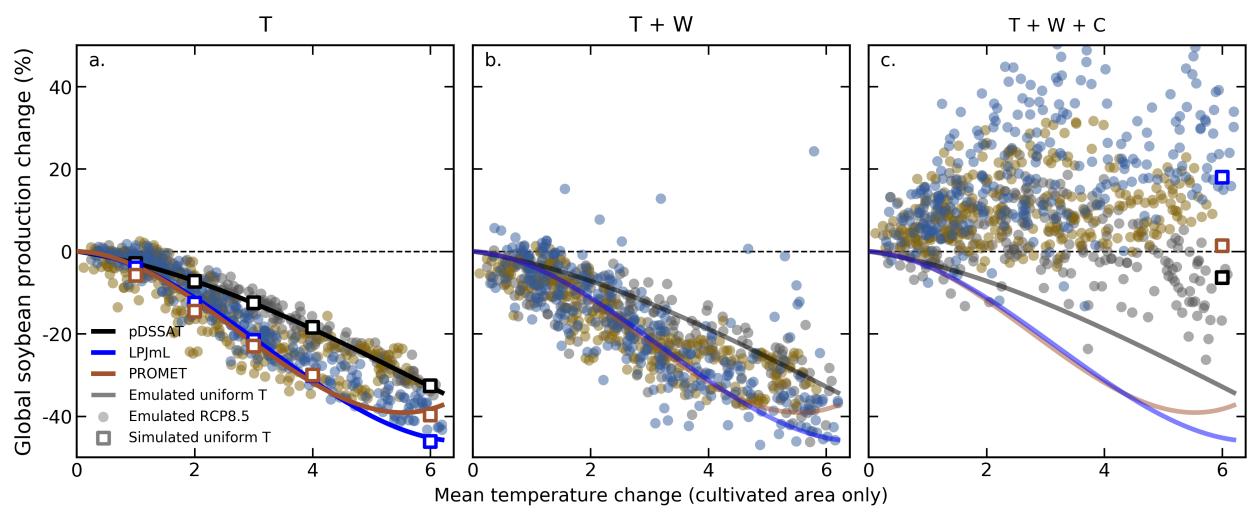


Figure S14: Illustration of the factors affecting yields in more realistic climate scenarios for soy. Same convention main text Figure 11.

7 Reduced specification (23-term) emulator examples

In this section we present some parallel figures for the reduced-form emulator (23-term). Issues with the reduced-form model are most prominent in PROMET for rice and soy, and JULES soy and spring wheat. Some potential causes behind this difference in ease of emulation are as follows. First, we do not emulate PROMET or JULES in the nitrogen dimension (CARAIB is the other model that does not do N). Both JULES and PROMET models are land system process models, originally developed with a broad focus, and have relatively recently (2015) been adapted for managed vegetation (i.e. agriculture). PROMET is the most sensitive model of all the models for rice in C, T, and W. PROMET is the quantitatively lowest performing model for soybeans when compared to the historical FAO data for the top 10 producing countries. JULES is the most sensitive model of all the models 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. CARAIB on the other hand (the other model that does not do nitrogen), was originally developed as a vegetation model in the early 90's and has a longer history of agricultural focus. It is the least sensitive model in T for rice and spring wheat, and is nearest to the ensemble mean in maize and winter wheat. In W, it is the least sensitive model for maize, and the wheats, the second least sensitive model for soybeans, and nearest to the ensemble mean for rice. For C, it is the nearest to the ensemble mean for the wheats and soy and unremarkable for rice. However, CARAIB is the most sensitive model for maize in the C dimension (though this is relatively less sensitive than most other responses for other models the other crops).

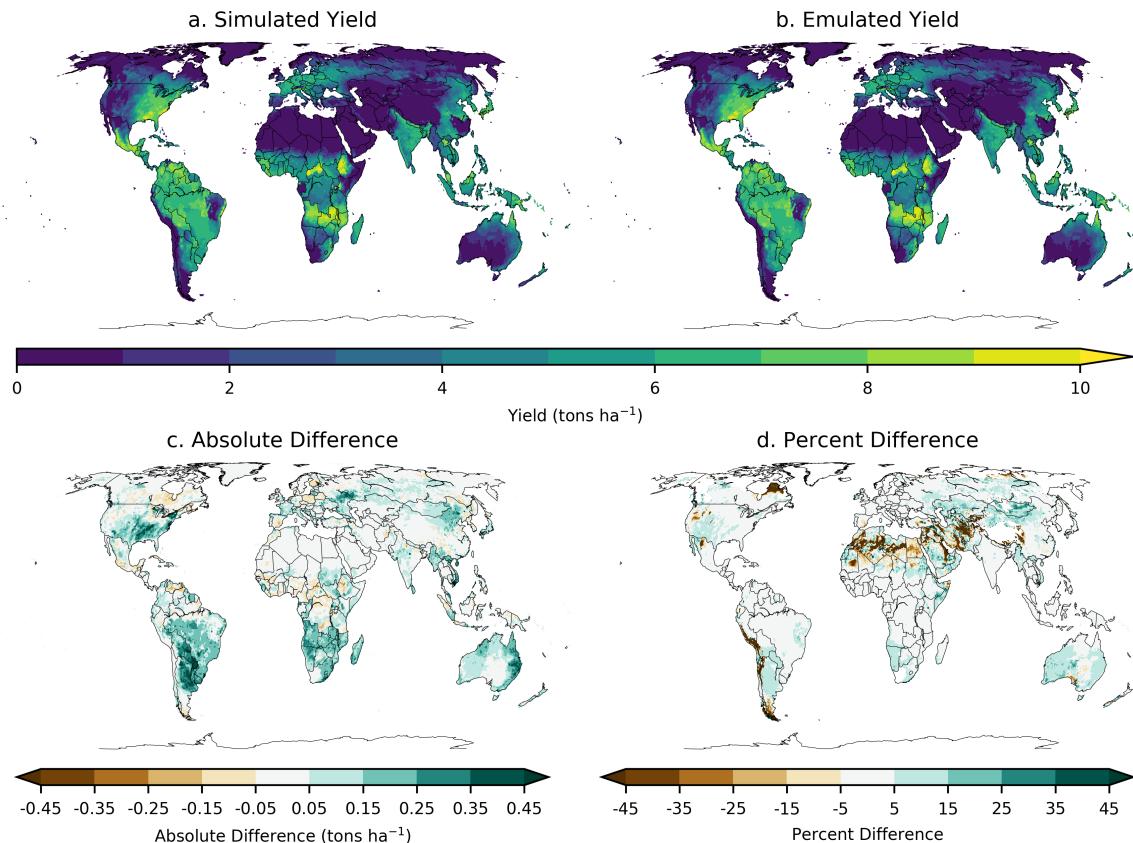


Figure S15: Same convention as Figure 4 in the main text except now with the reduced (23-term) emulator specification.

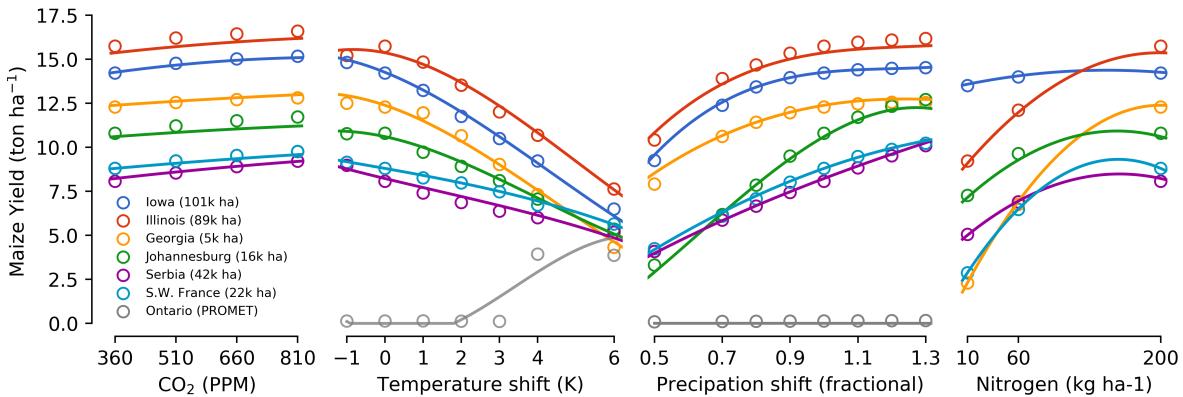


Figure S16: Same convention as Figure 5 in the main text except now with the reduced (23-term) emulator specification.

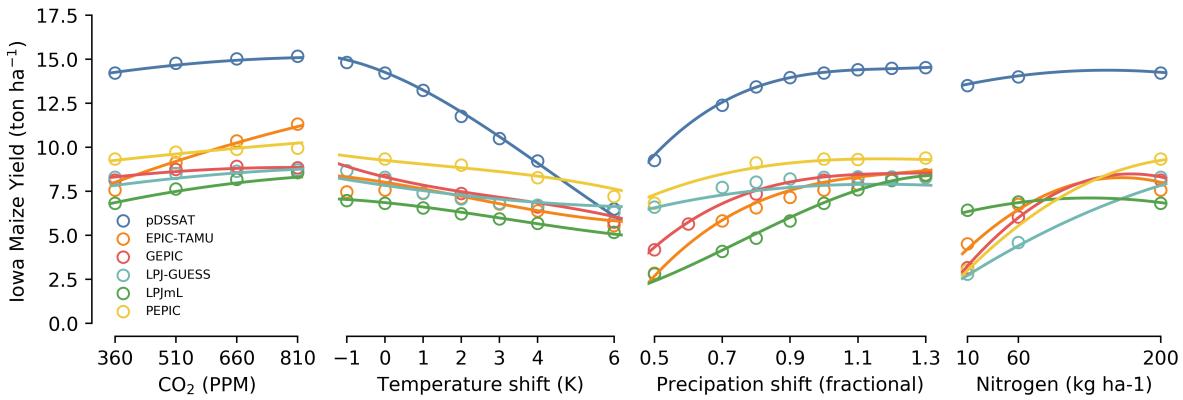


Figure S17: Same convention as Figure 6 in the main text except now with the reduced (23-term) emulator specification.

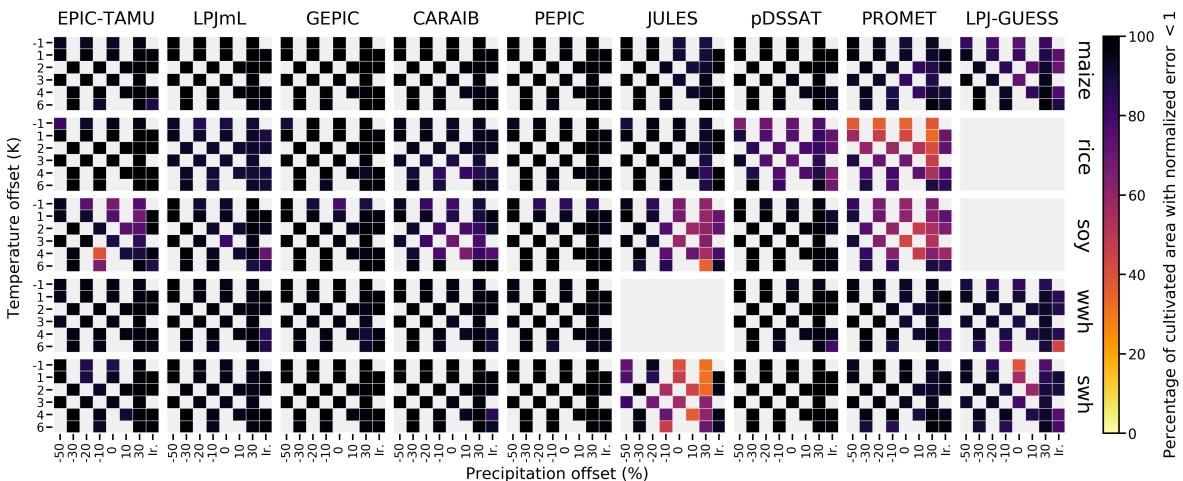


Figure S18: Same convention as Figure 7 in the main text except now with the reduced (23-term) emulator specification.

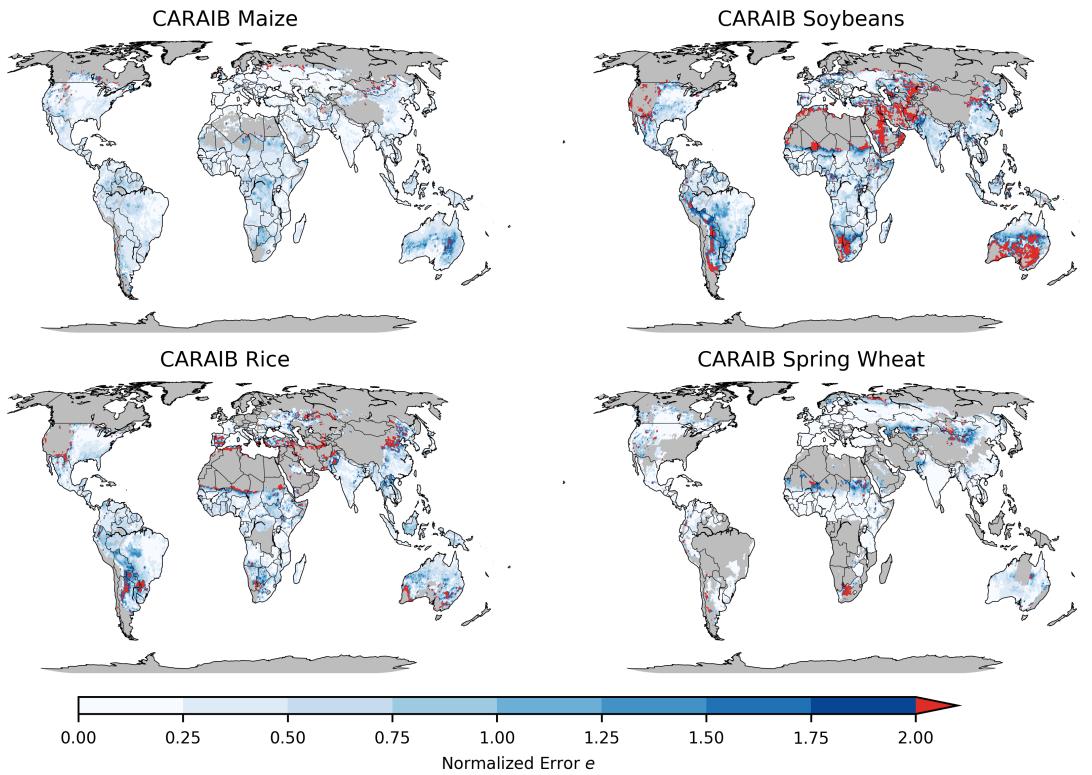


Figure S19: Same convention as Figure 8 in the main text except now with the reduced (23-term) emulator specification.

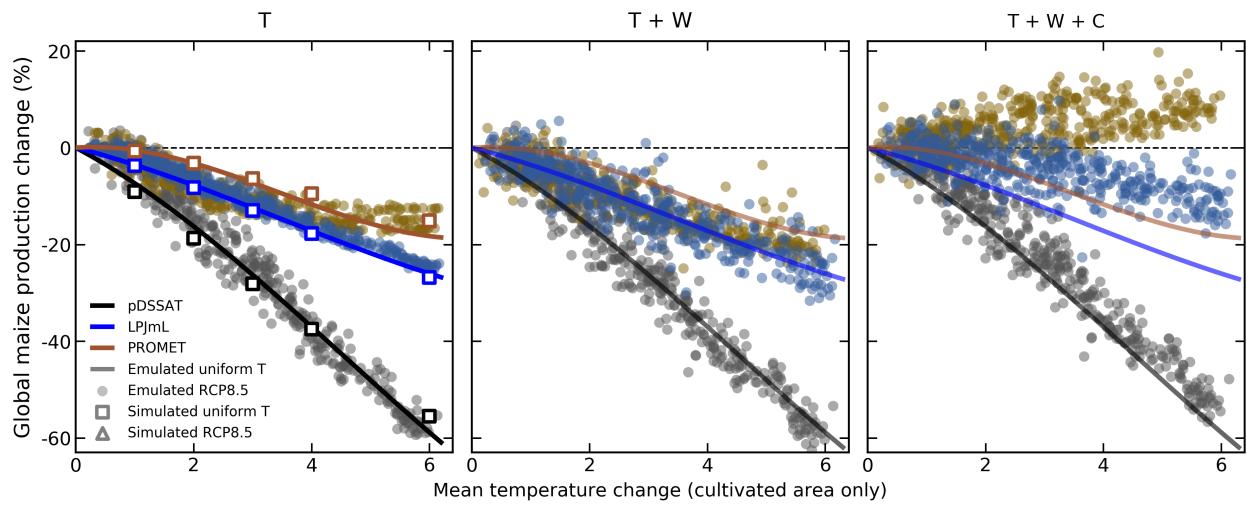


Figure S20: Same convention as Figure 11 in the main text except now with the reduced (23-term) emulator specification.

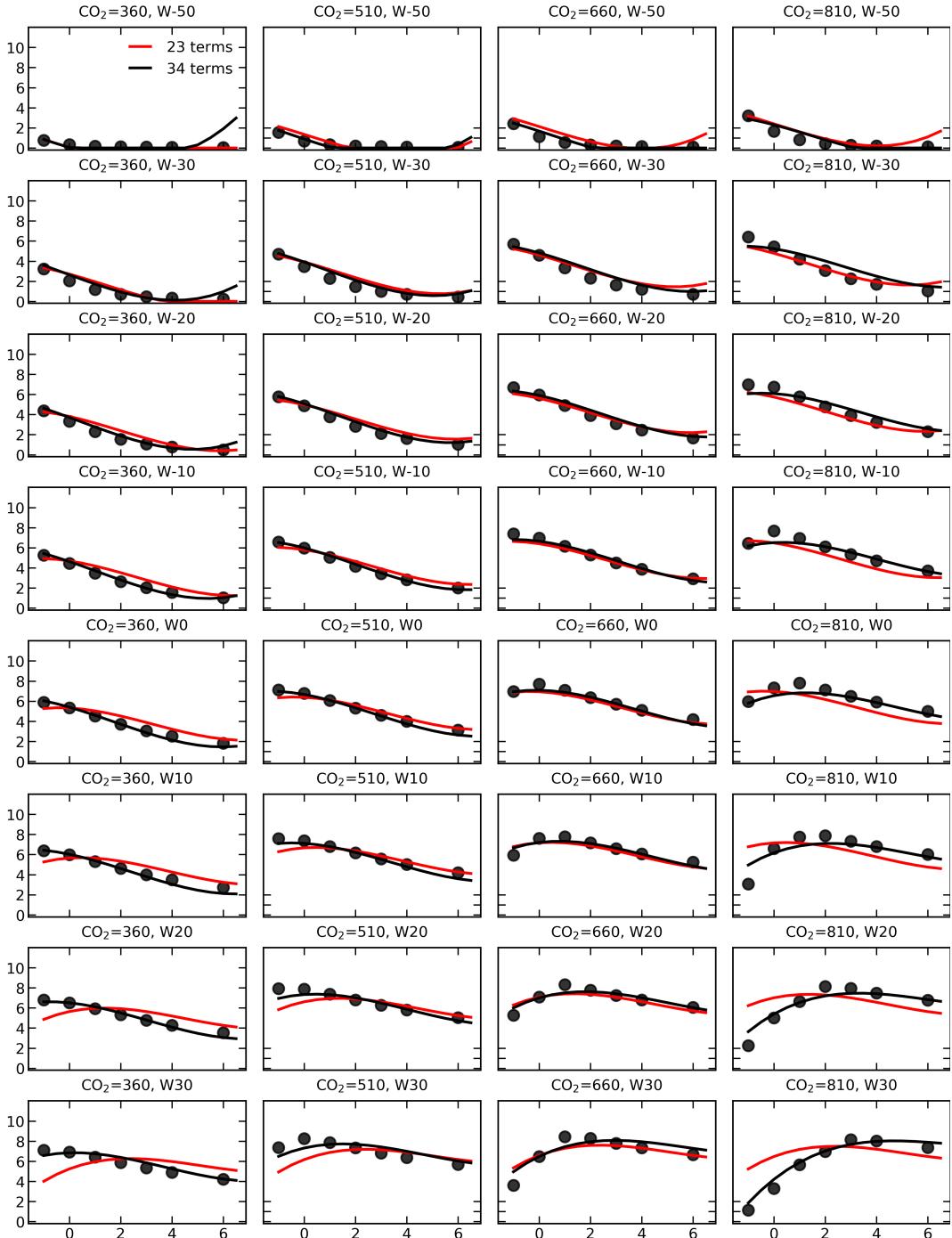


Figure S21: Example of emulator failure. Simulated and emulated values for JULES soybean in Southern Germany. RMSE = 41% of baseline yield for the reduced form (23-term) emulator. The downturn in yields as C and W increase can only be captured by the higher order C interaction terms.

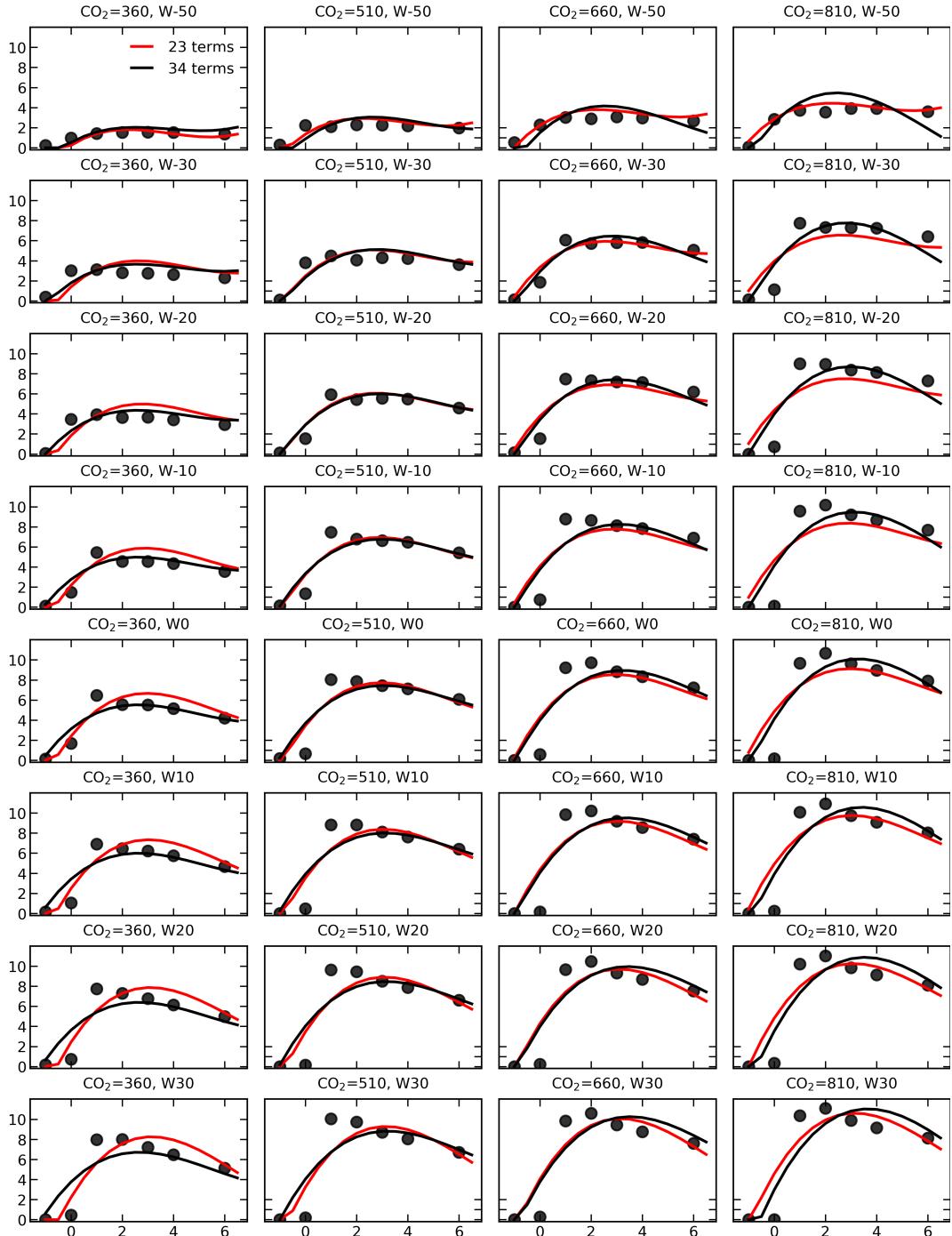


Figure S22: Example of emulator failure. Simulated and emulated values for PROMET rice in Arunachal Pradesh. RMSE = 132% of baseline yield for red (reduced fit). The step change in the yields around 0 K at higher water specifications cannot be captured by any third order polynomial. Both emulator specifications fail in this example.

8 Yield Responses for other crops and models

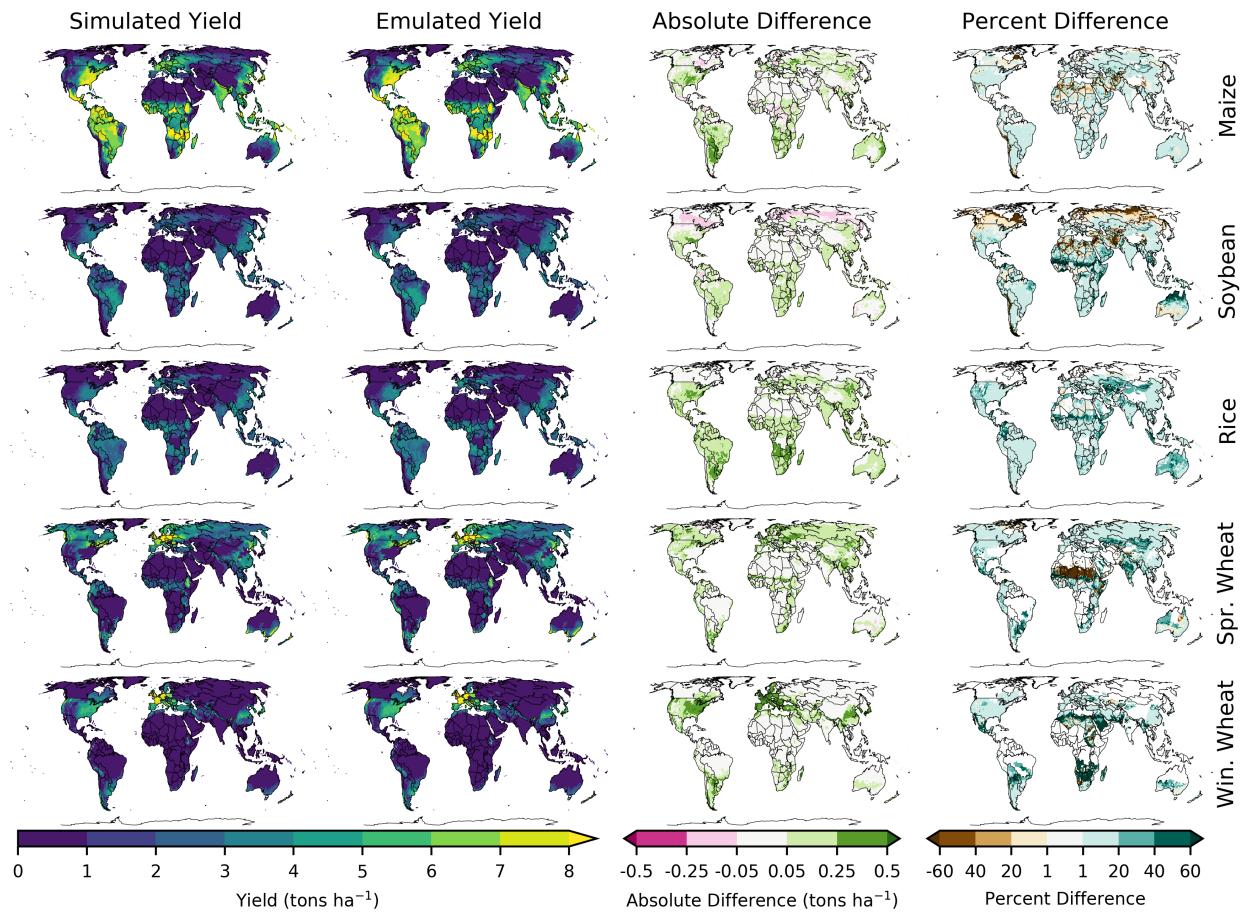


Figure S23: Spatial yield response and emulator error for LPJmL for all crops included in the stud. Same convention as Figure 4 in the main text.

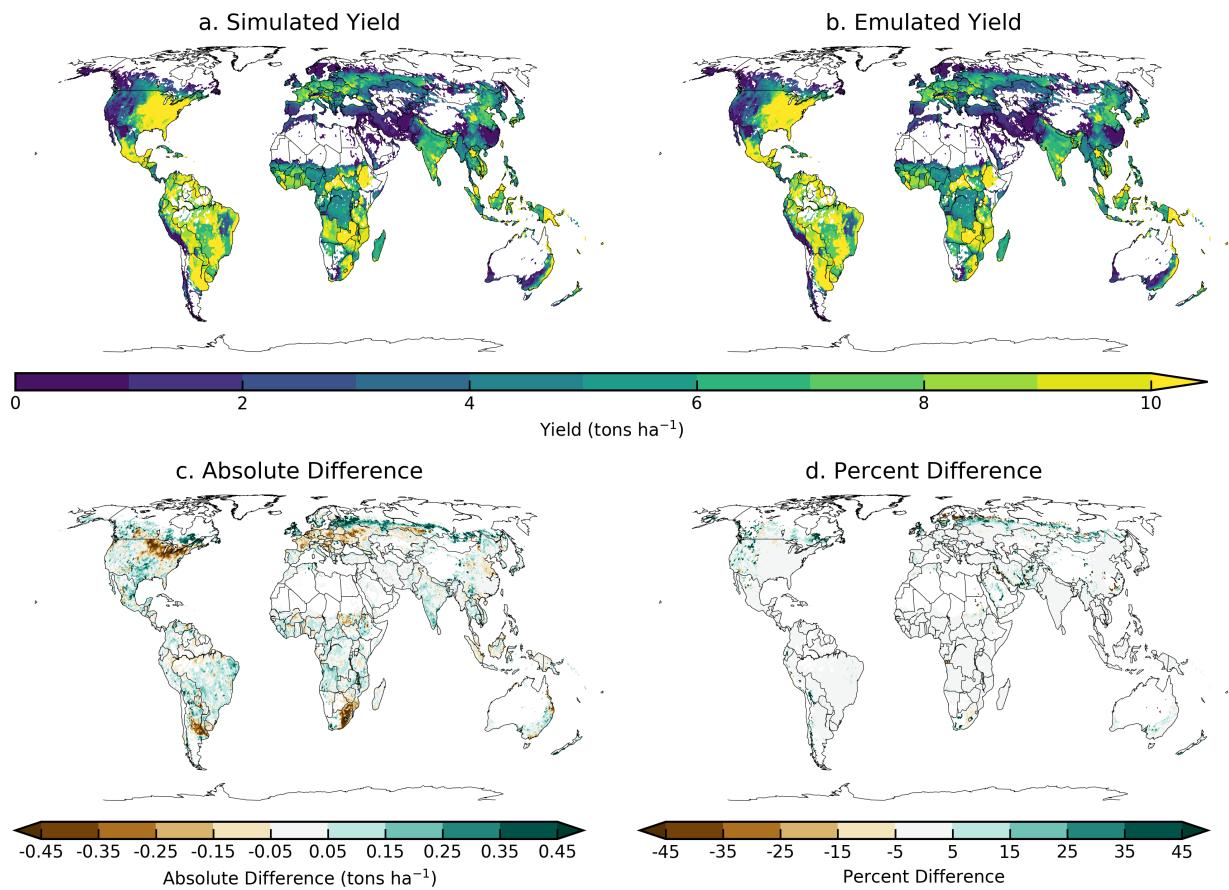


Figure S24: Spatial yield response and emulator error for pDSSAT for maize. Same convention as Figure 4 in the main text.

9 Normalized Error for all models

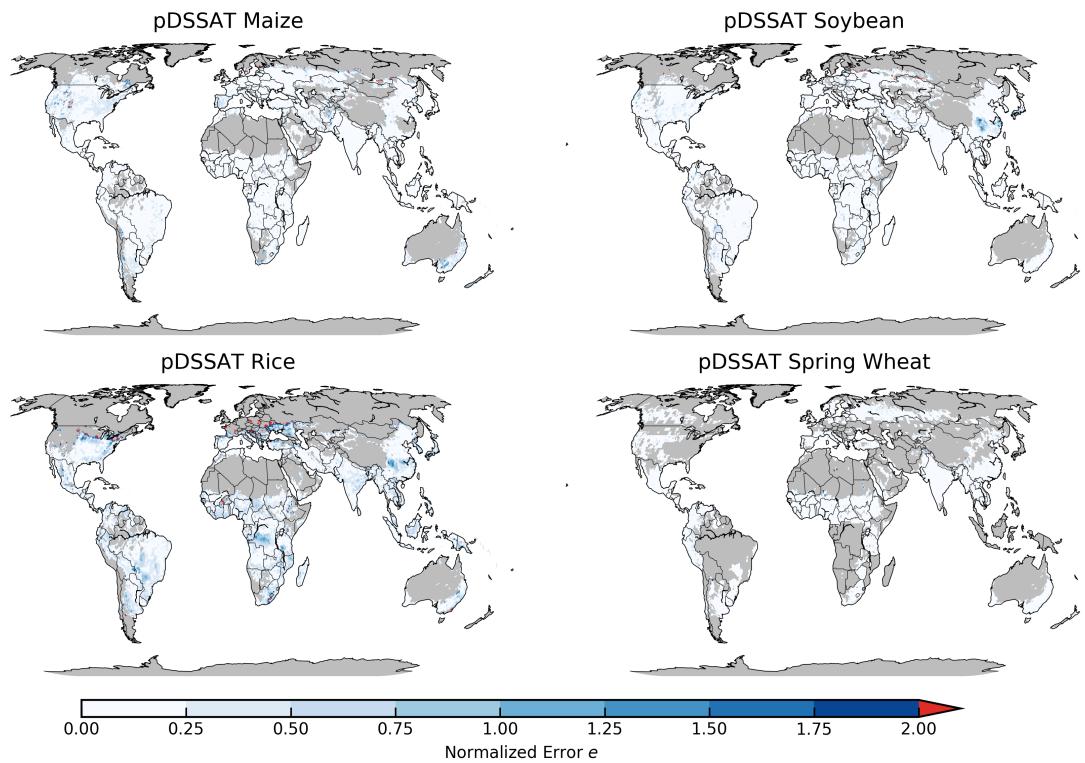


Figure S25: Normalized error e for pDSSAT. Same convention as main text Figure 8.

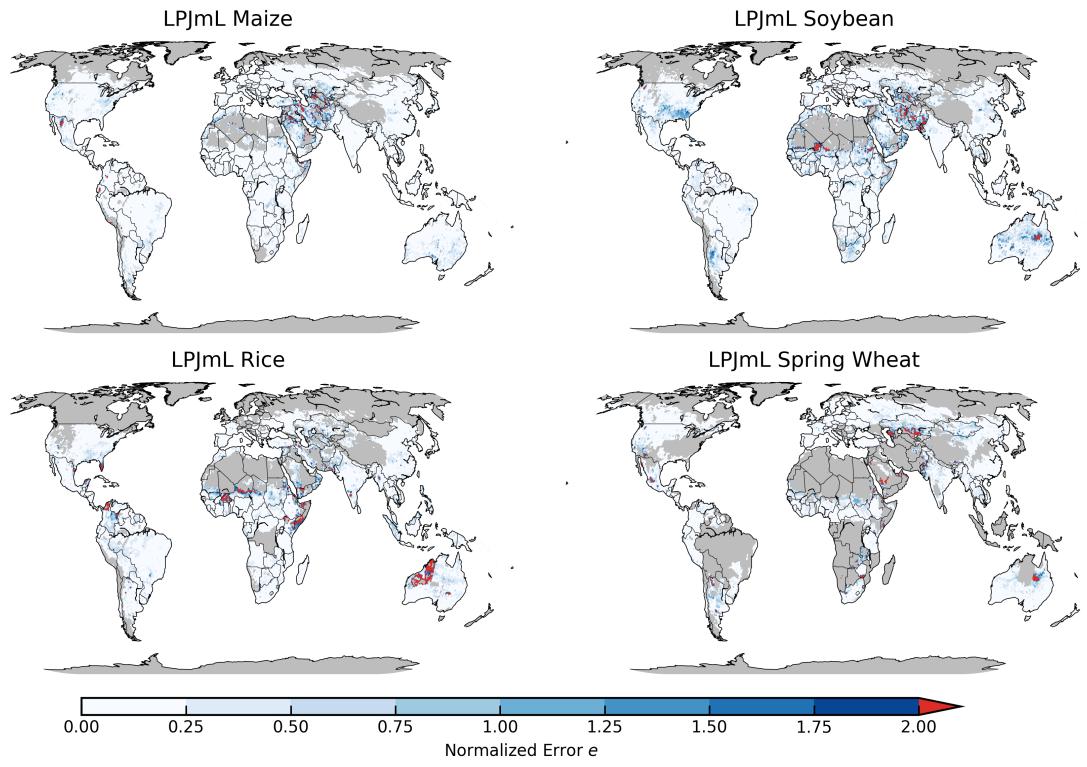


Figure S26: Normalized error e for LPJmL. Same convention as main text Figure 8.

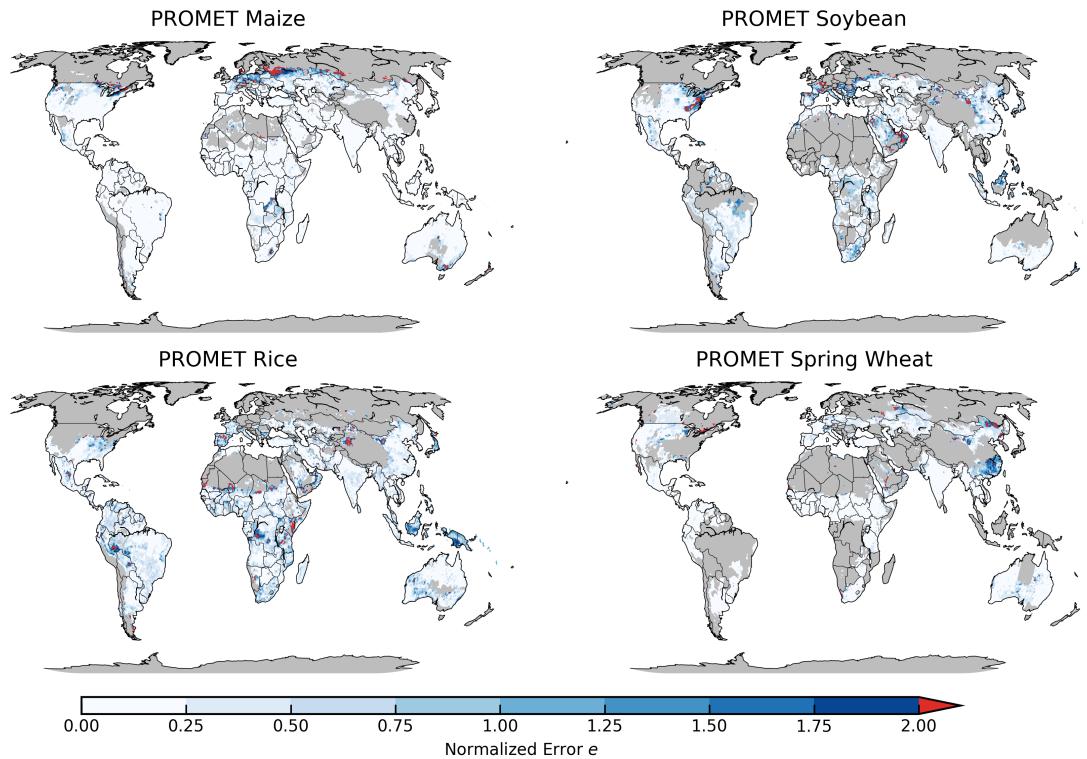


Figure S27: Normalized error e for PROMET. Same convention as main text Figure 8.

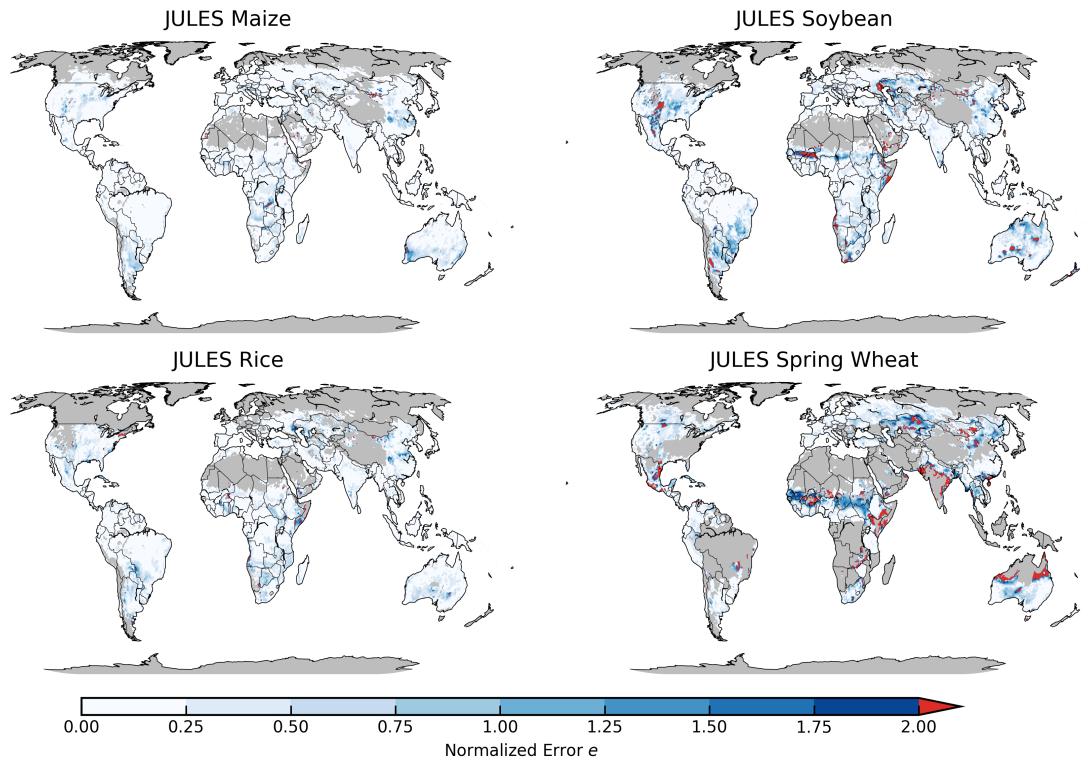


Figure S28: Normalized error e for JULES. Same convention as main text Figure 8.

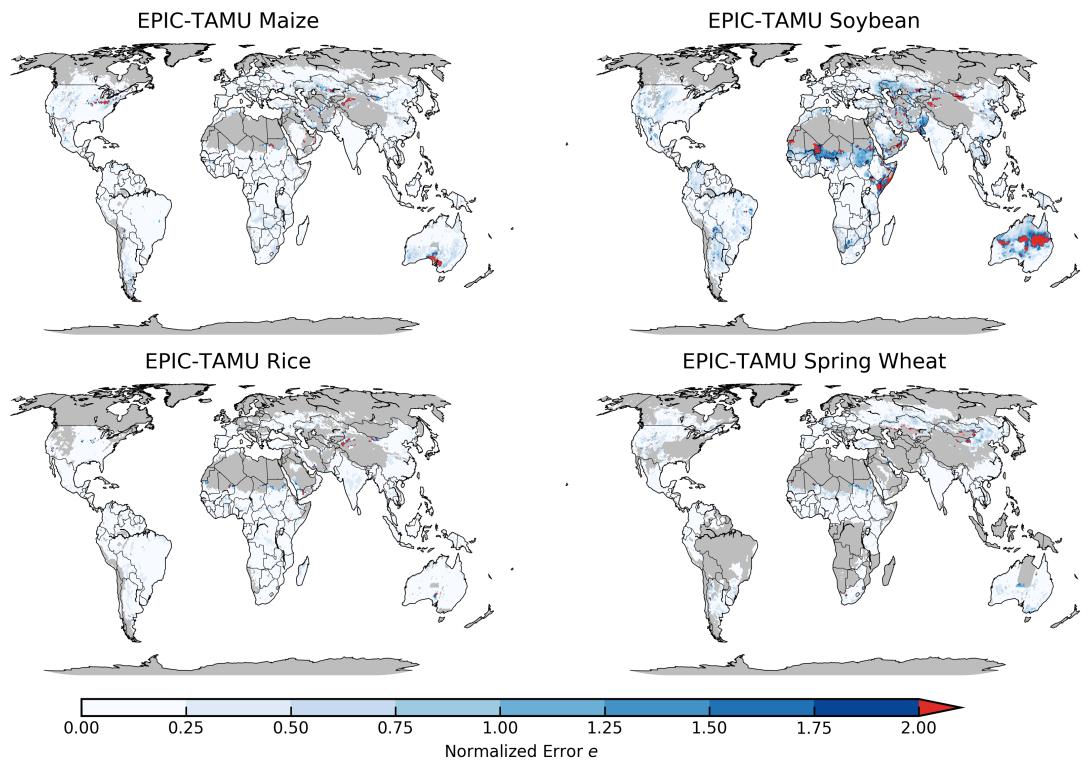


Figure S29: Normalized error e for EPIC-TAMU. Same convention as main text Figure 8.

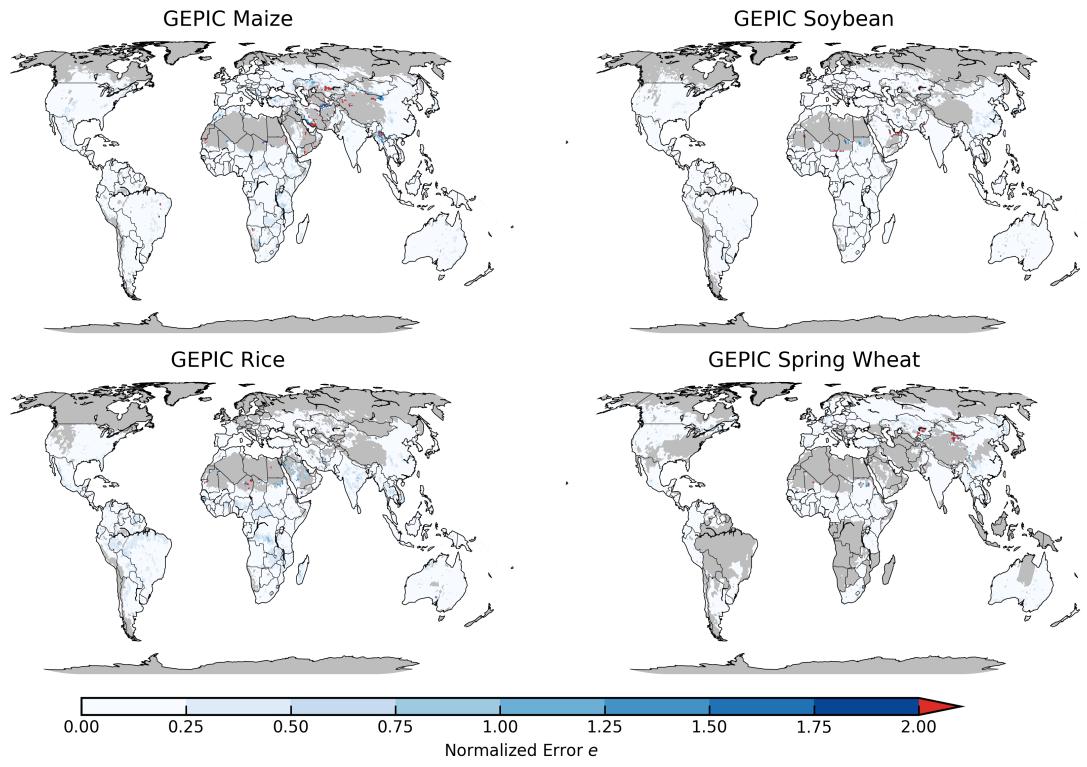


Figure S30: Normalized error e for GEPIC. Same convention as main text Figure 8.

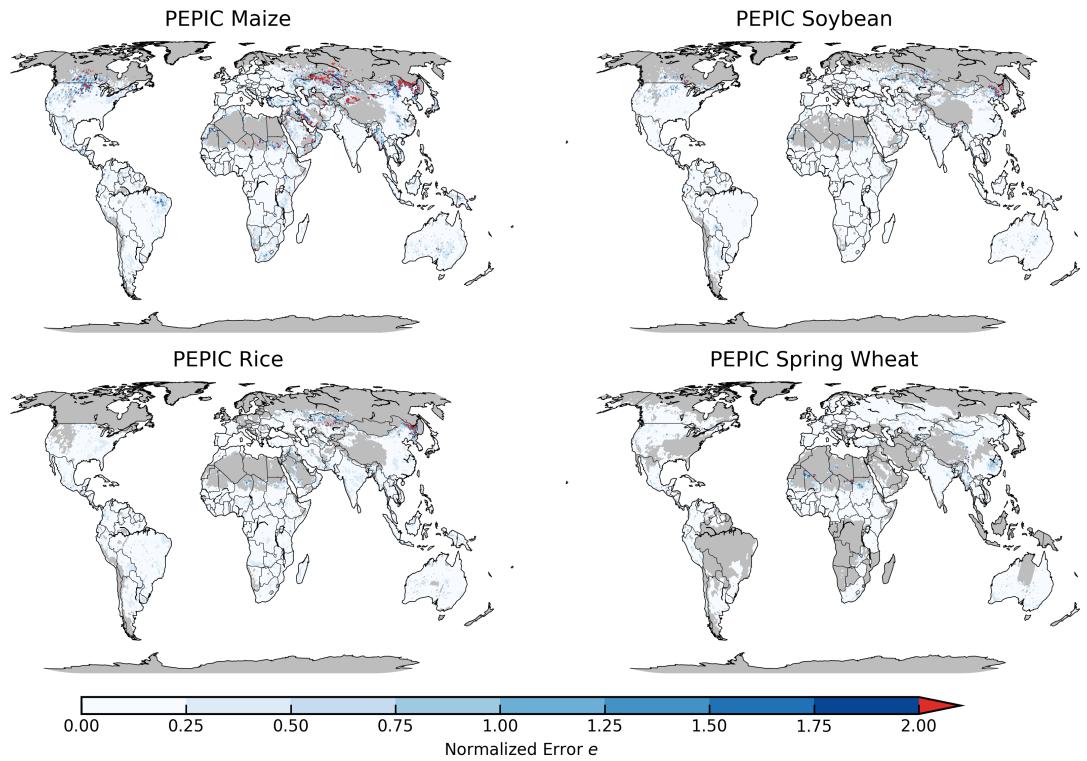


Figure S31: Normalized error e for PEPIC. Same convention as main text Figure 8.

10 Cross validation error for all models

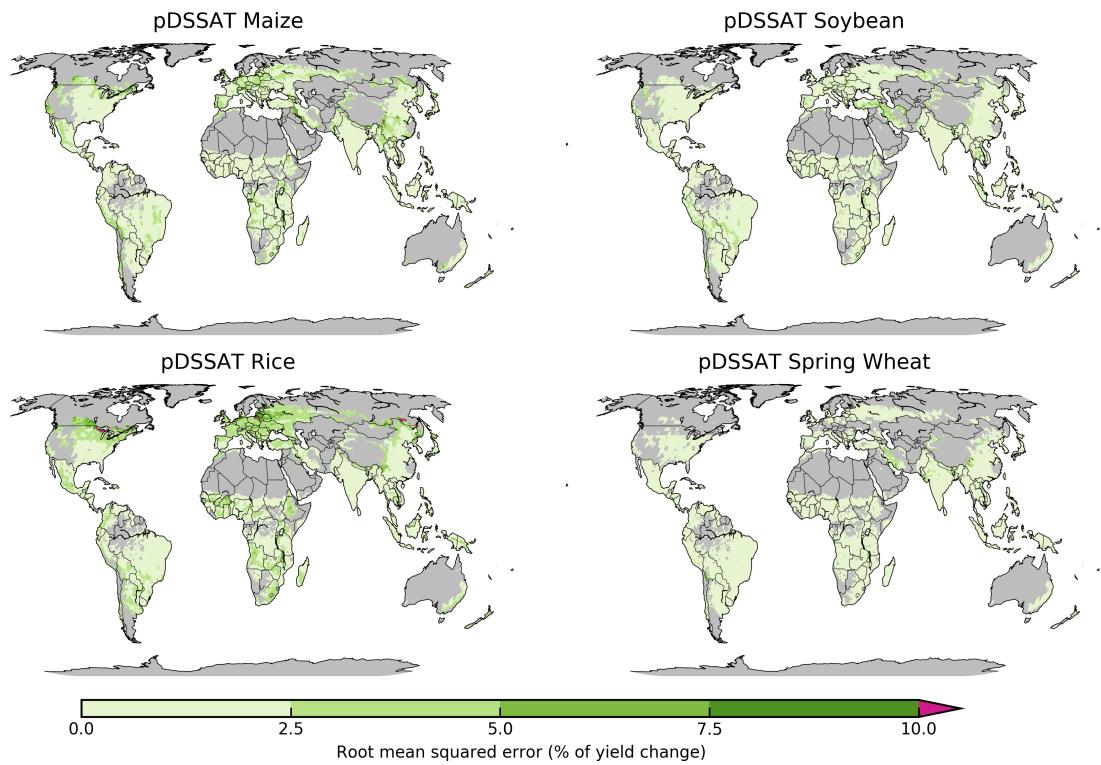


Figure S32: Root mean squared error for three-fold cross validation for the pDSSAT model for rainfed crops. Values shown as a percentage of baseline yield in each gridcell.

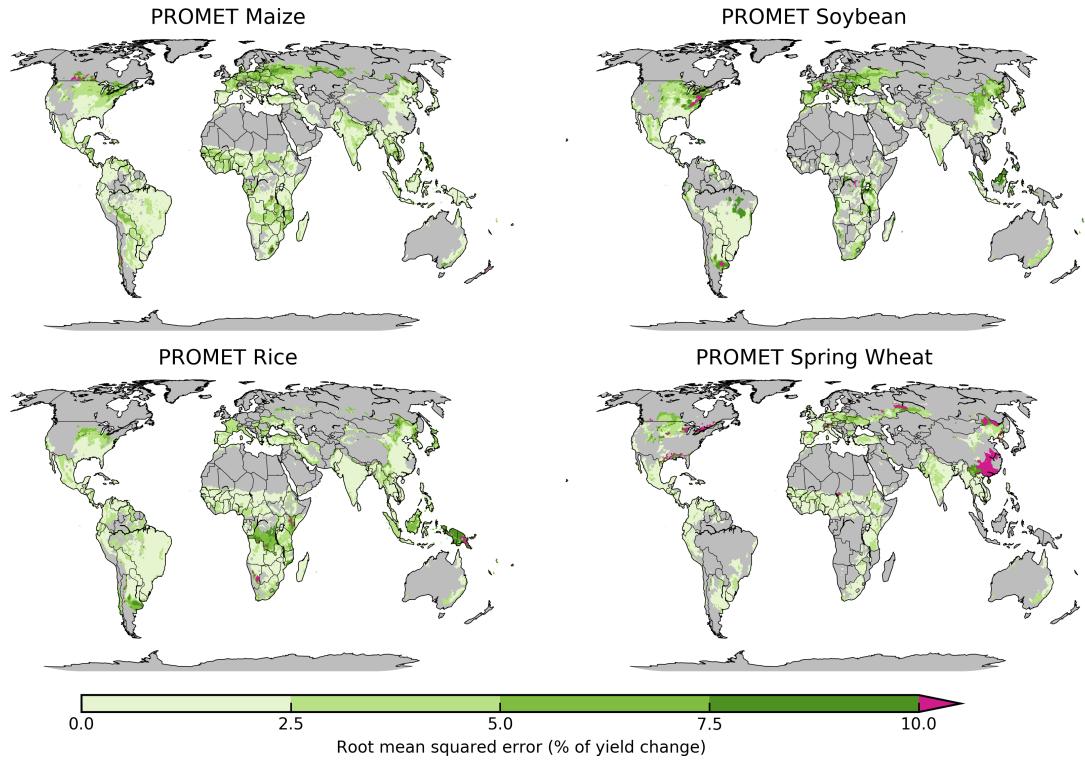


Figure S33: Map of root mean squared error for three fold cross validation process for the PROMET model for rainfed crops. Values shown as a percentage of baseline yield in each gridcell.

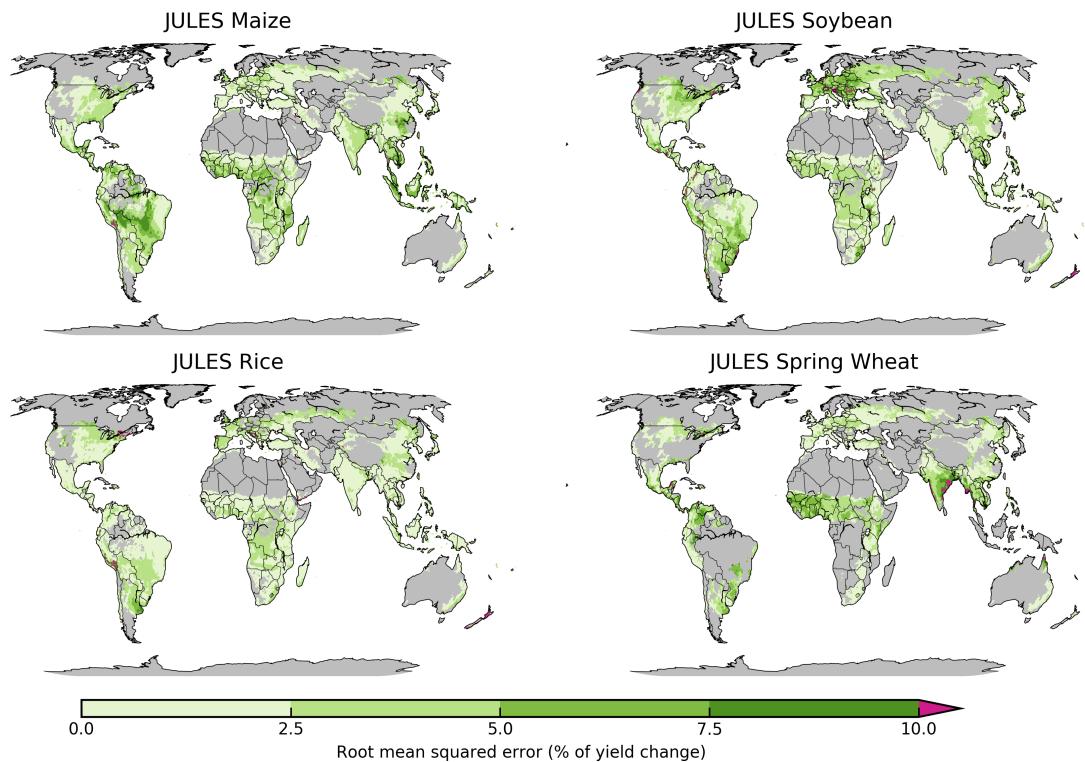


Figure S34: Map of root mean squared error for three fold cross validation process for the JULES model for rainfed crops. Values shown as a percentage of baseline yield in each gridcell.

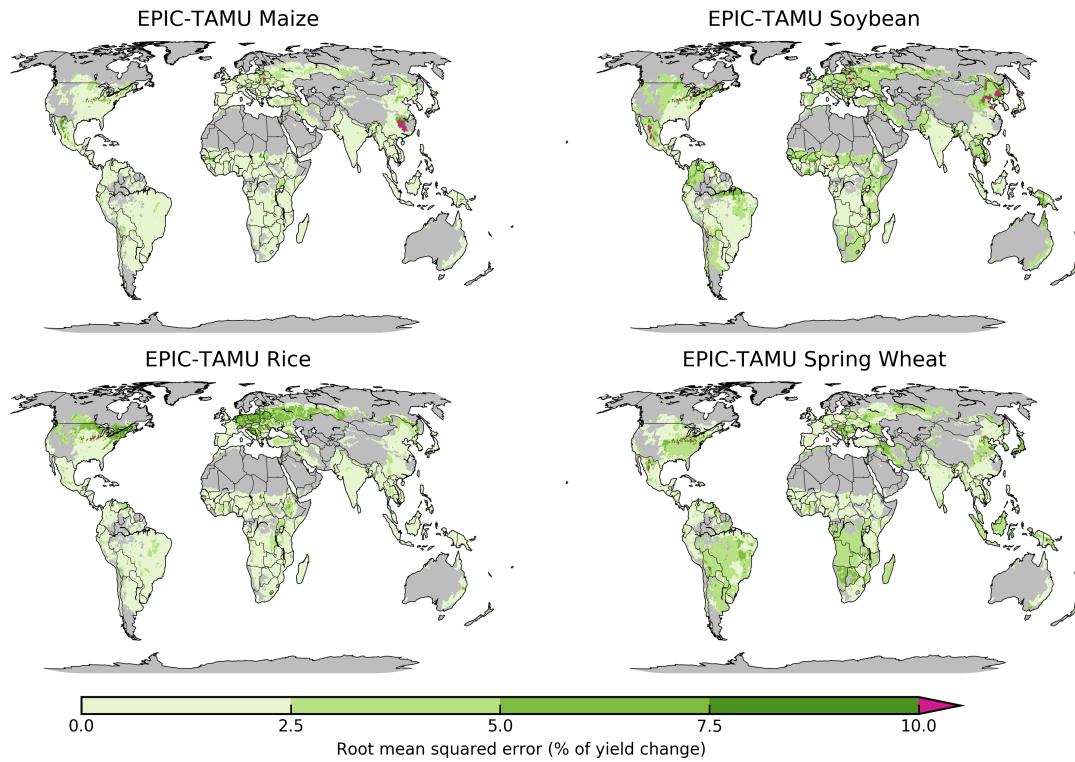


Figure S35: Map of root mean squared error for three fold cross validation process for the EPIC-TAMU model for rainfed crops. Values shown as a percentage of baseline yield in each gridcell.

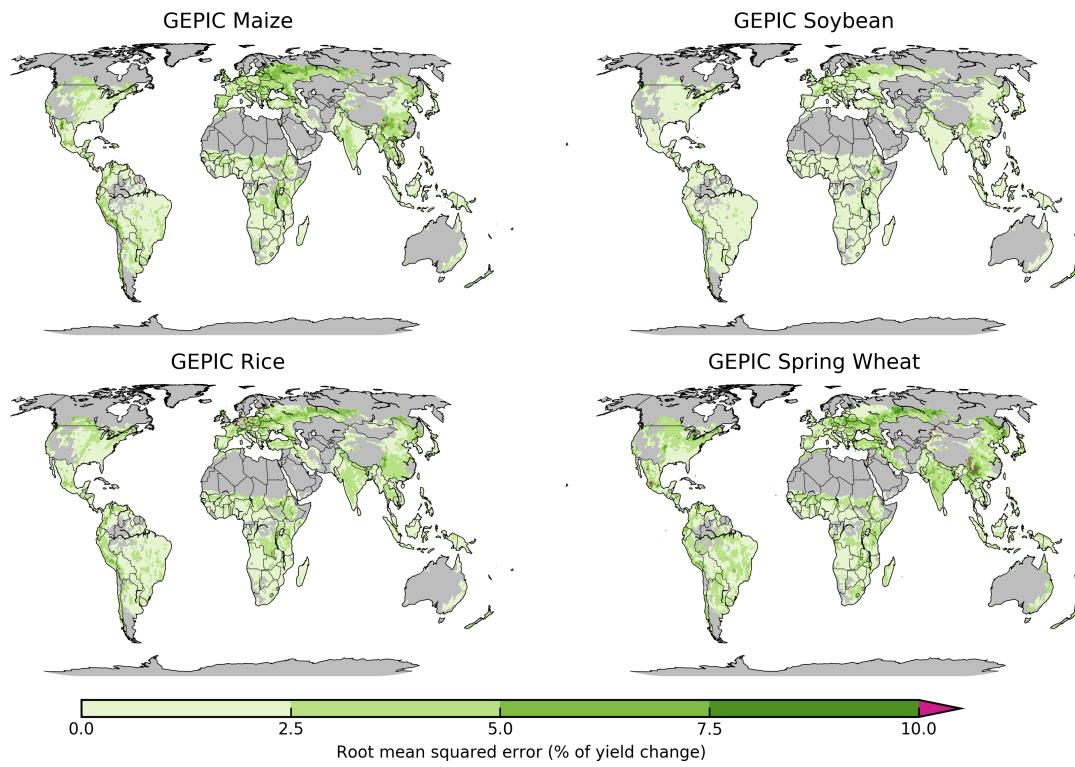


Figure S36: Map of root mean squared error for three fold cross validation process for the GEPIC model for rainfed crops. Values shown as a percentage of baseline yield in each gridcell.

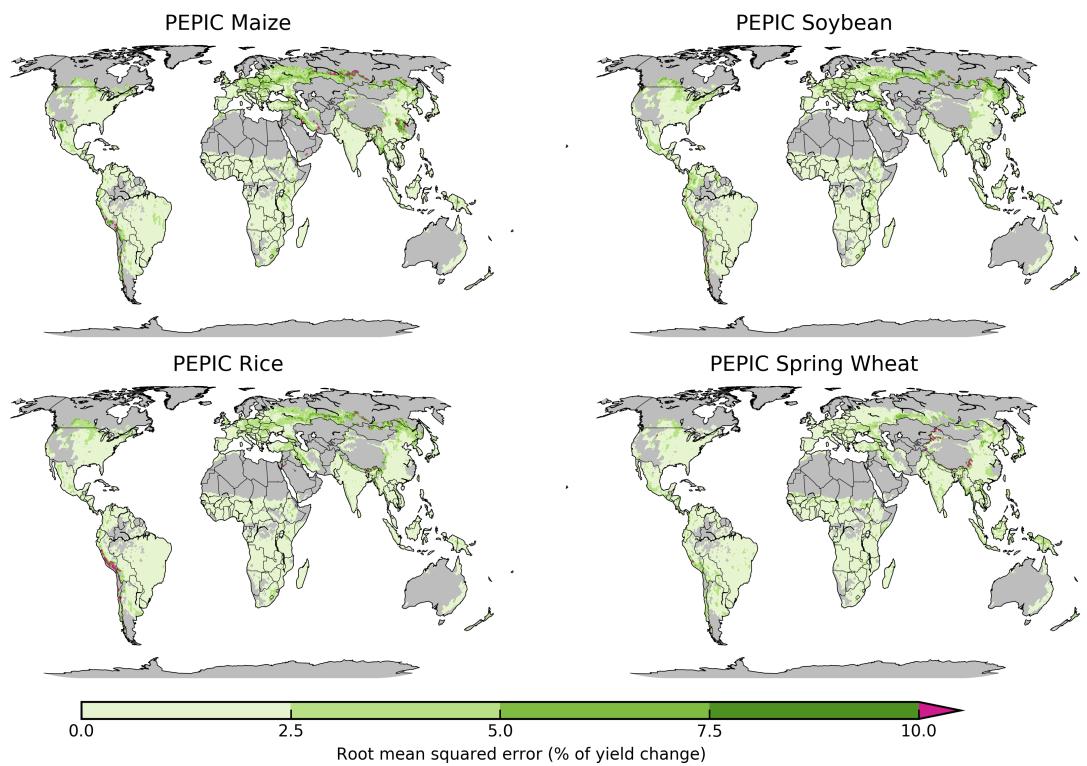


Figure S37: Map of root mean squared error for three fold cross validation process for the PEPIC model for rainfed crops. Values shown as a percentage of baseline yield in each gridcell.