

**Supplemental Material**  
**The GGCMI Phase II experiment: simulating and emulating  
global crop yield responses to changes in carbon dioxide,  
temperature, water, and nitrogen levels**

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

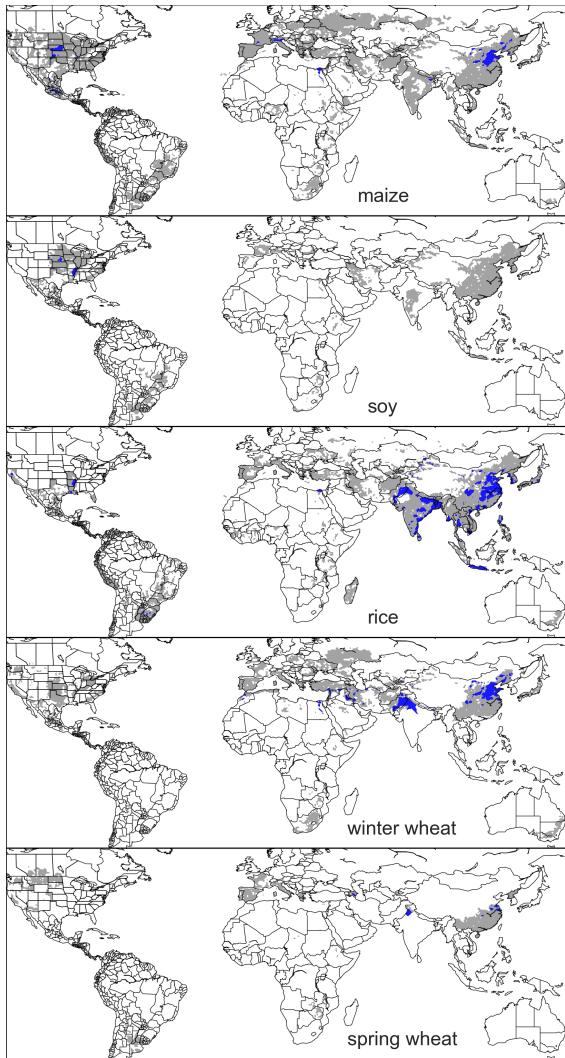
1. Department of the Geophysical Sciences, University of Chicago, Chicago, IL, USA
2. Center for Robust Decision-making on Climate and Energy Policy, University of Chicago, Chicago, IL, USA
3. Department of Computer Science, University of Chicago, Chicago, IL, USA
4. Potsdam Institute for Climate Impact Research, Leibniz Association (Member), Potsdam, Germany
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. Ecosystem Services and Mgm. Prg., International Institute for Applied Systems Analysis, Laxenburg, Austria
8. Department of Soil Science, Comenius University in Bratislava, Bratislava, Slovak Republic
9. Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, 91191 Gif-sur-Yvette, France
10. Sino-French Institute of Earth System Sciences, Peking University, Beijing, China
11. Unité de Modélisation du Climat et des Cycles Biogéochimiques, University of Liège, Belgium
12. Met Office Hadley Centre, Exeter, United Kingdom
13. Department of Geography, Ludwig-Maximilians-Universität, Munich, Germany
14. Georg-August-University Göttingen, Tropical Plant Production and Ag. Sys. Modelling, Göttingen, Germany
15. Department of Geographical Sciences, University of Maryland, College Park, MD, USA
16. Texas Agrilife Research and Extension, Texas A&M University, Temple, TX, USA
17. Department of Statistics, University of Chicago, Chicago, IL, USA
18. EAWAG, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland
19. Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden
20. Earth Institute Center for Climate Systems Research, Columbia University, New York, NY, USA
21. Karlsruhe Institute of Technology, IMK-IFU, 82467 Garmisch-Partenkirchen, Germany
22. School of Geography, Earth and Environmental Science, University of Birmingham, Birmingham, UK

## Contents

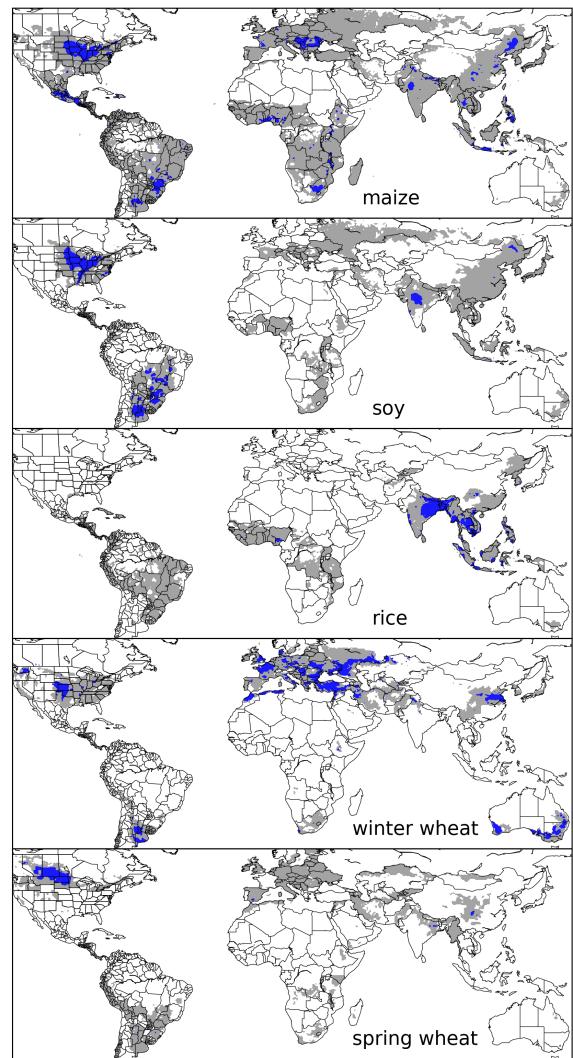
<b>1 Methods</b>	<b>2</b>
<b>2 Simulation results</b>	<b>5</b>
<b>3 Emulator fits and performance</b>	<b>9</b>
<b>4 Emulator results</b>	<b>12</b>

# 1 Methods

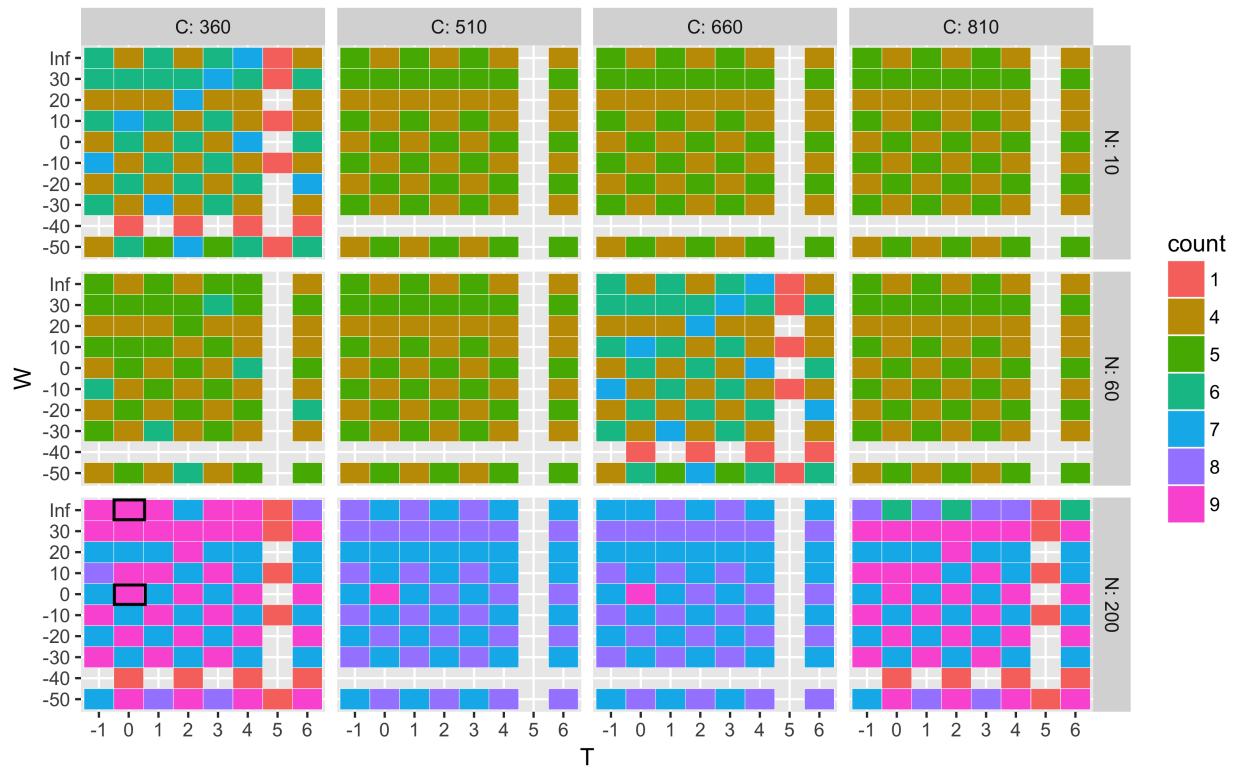
- 2 Additional example figures for reference in the methods section. As noted in the manuscript. Manuscript figure also repeated for ease of comparison.



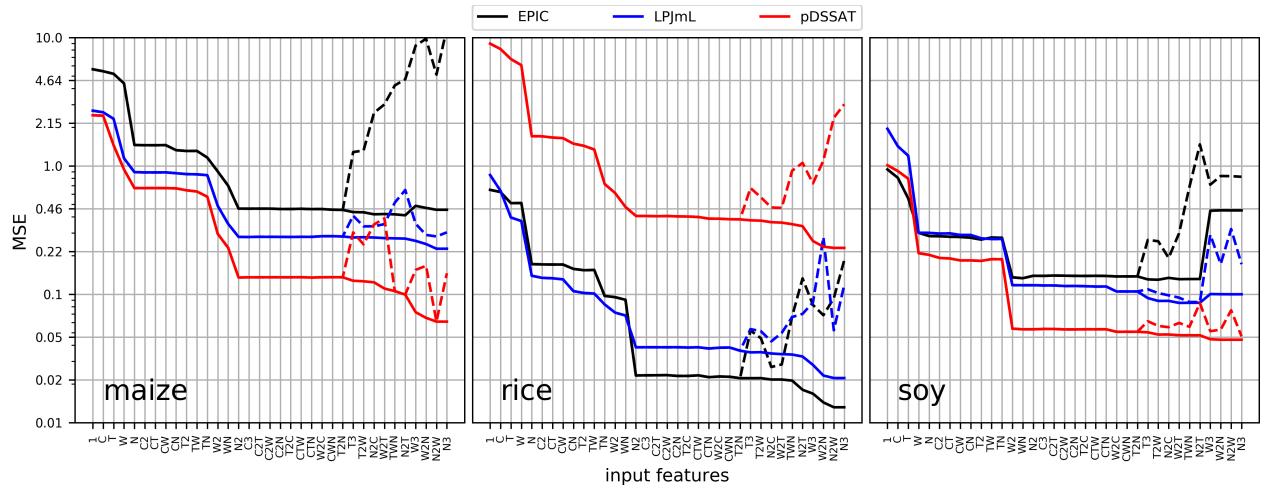
**Figure S1:** Presently cultivated area for irrigated crops in the real world. The blue contour area indicates grid-cells with more than 20,000 hectares of crop cultivated. The gray contour shows area with more than 10 hectares cultivated. Data from the MIRCA2000 data set for maize, rice, and soy. Winter and spring wheat areas are adapted from MIRCA2000 data and sorted by growing season.



**Figure S2:** Presently cultivated area for rain fed crops in the real world. Conventions as in Figure S1. This figure repeats manuscript Figure 1.



**Figure S3:** Tile 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. Error calculations are run over scenarios where all models provided a simulation. Three models were excluded from the emulator analysis due to low sampling. Error calculations are run over scenarios with max number of models (See Figures S12, S13).

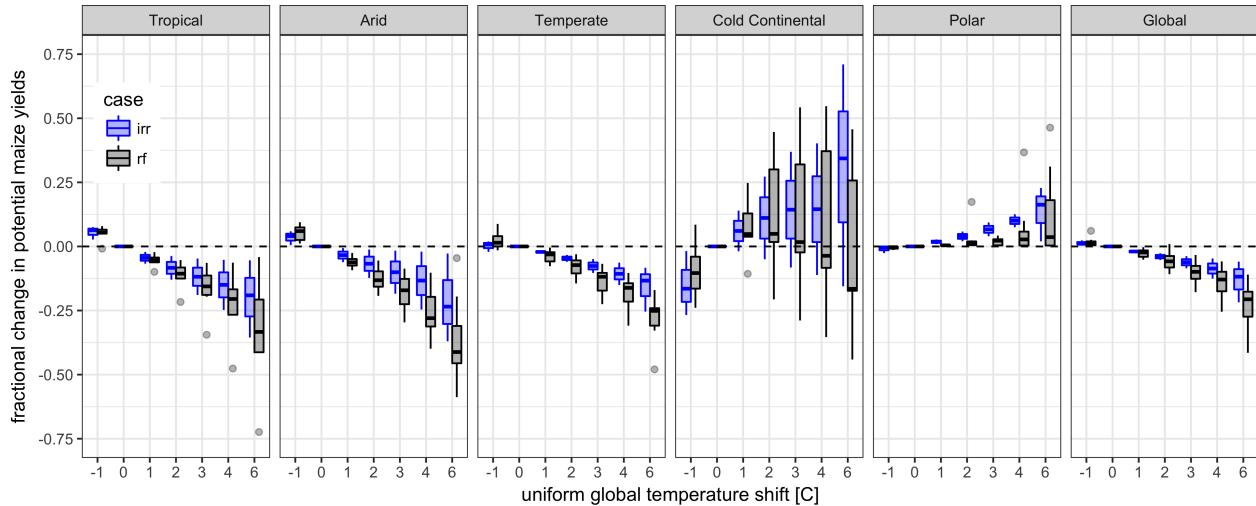


**Figure S4:** Mean Squared Error (MSE) from the cross-validation process. The emulate yield for maize rice and soy are compared to the simulated values for the same model in all grid cells where crops are actually cultivated, weighted by area in each grid cell. X- axis indicates terms included in the model at each step progressively where T = temperature and T2 =Terms that did not reduce the aggregate error are not included in the final model. Some terms that did not reduce the aggregate error are included if a higher order version of that term provided a decrease in mean squared error (i.e. temperature cubed cannot be included without the temperature squared term, and the linear temperature term). Even though the models exhibit different absolute levels of error, all three models agree remarkably well on feature importance. Solid lines indicate L1 normalization and dashed lines indicate L2 normalization of yield outputs. Colors indicated different models (three fully-sampled simulation sets.)

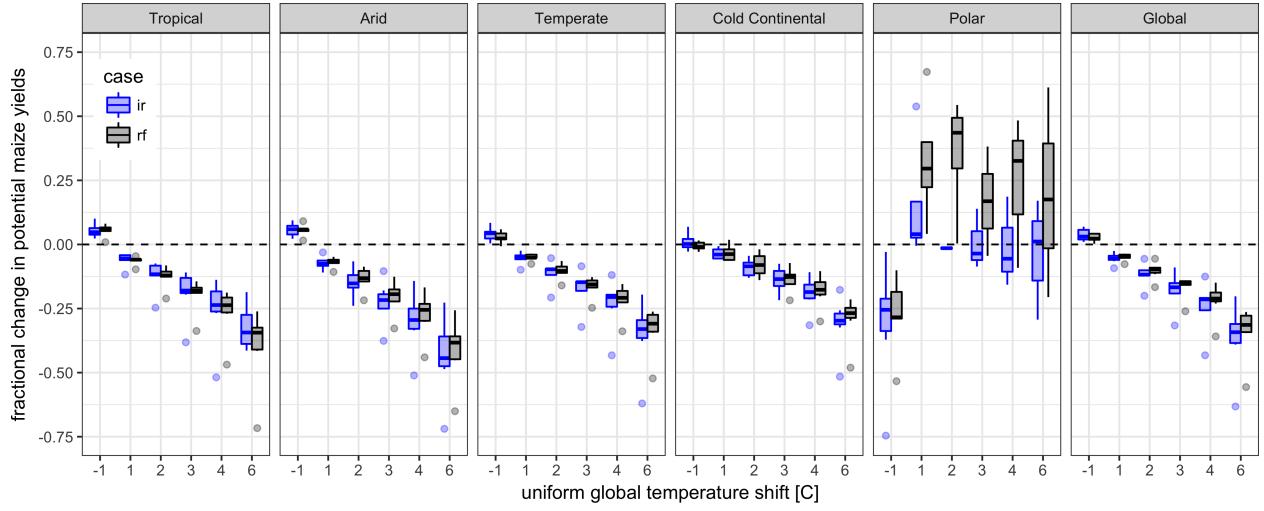
## 2 Simulation results

We present additional simulation results for illustration as noted in the manuscript. Most crops exhibit a somewhat uniform response to temperature increase across different Köppen-Geiger when analyzed over currently cultivated area (see Figure S16: i.e. equatorial maize and ‘snow’ maize show similar response to a temperature increase). This counterintuitive result agrees with existing literature including ? which shows increases in yields mainly in regions where crops are not currently grown and in ?. A primary cause of this effect is less difference in growing season temperature across Köppen-Geiger regions when they are weighted by current cultivation area than might be expected. Additionally, it has been proposed that the growing season is shortened under warmer temperatures in a way that is independent of baseline growing season temperature (e.g. ??). Currently most models in GGCMI include a direct linear shortening of the growing season with warming, but uncertainty about the exact nature of this response remains and it is an active area of research.

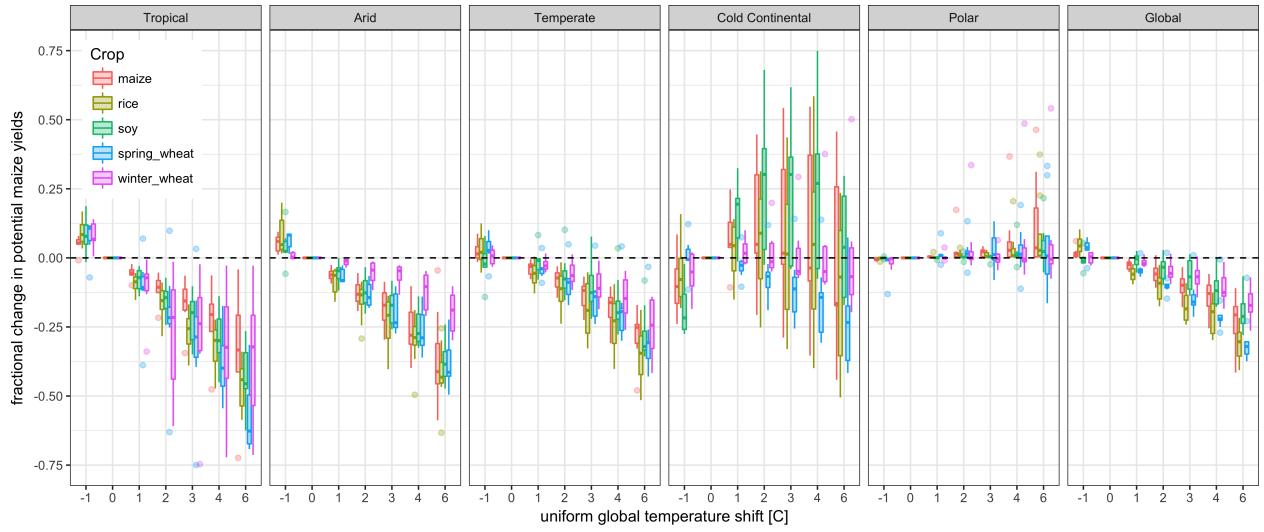
The CO<sub>2</sub> response is generally subject to large uncertainties (not evident in Figures ?? – ?? for maize as it is a C4 crop). All relevant CO<sub>2</sub> processes have not been studied in sufficient detail or have not been implemented in models sufficiently (e.g. ?) and a broader experimental basis for model parameterization is needed (?)



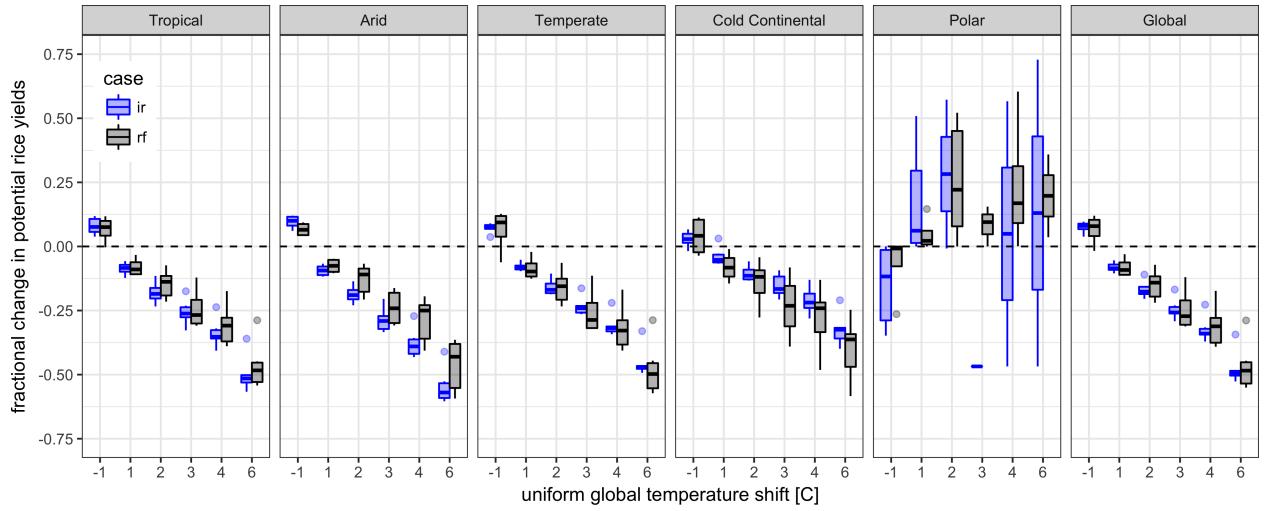
**Figure S5:** Irrigated maize has similar inter model variance as the baseline rain-fed case, but, as expected, shows greater resilience to temperature increases (reduced yield losses), especially in arid regions. All other covariates are held constant.



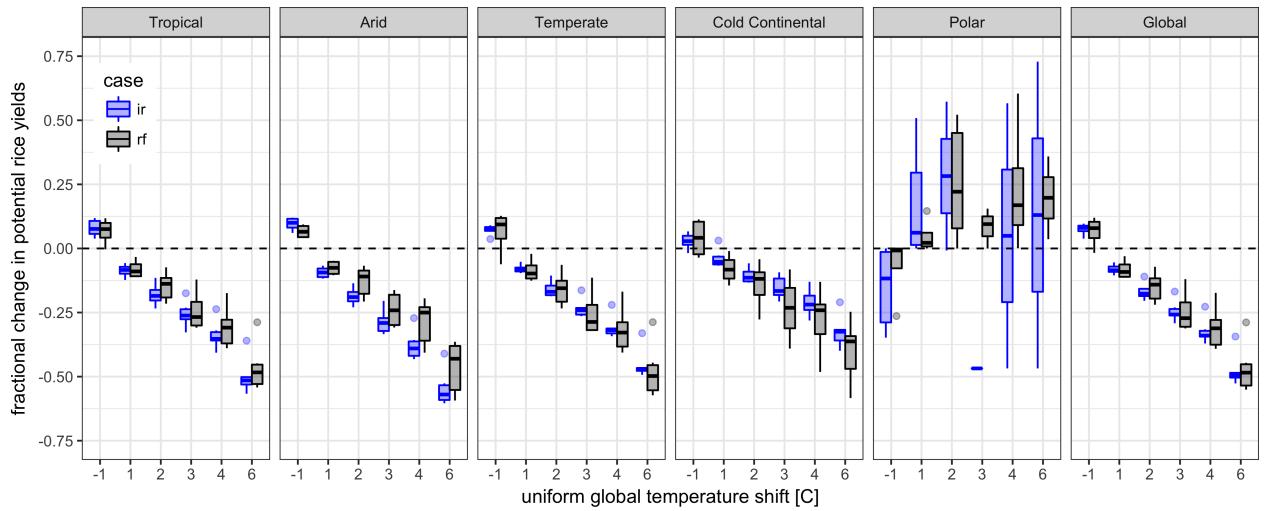
**Figure S6:** Same figure as above but weighted by actual cultivation area in the real world. Due to difference in cultivation area, irrigated maize is more sensitive to increased temperature on currently cultivated area. All other covariates are held constant. PROMET and JULES not shown.



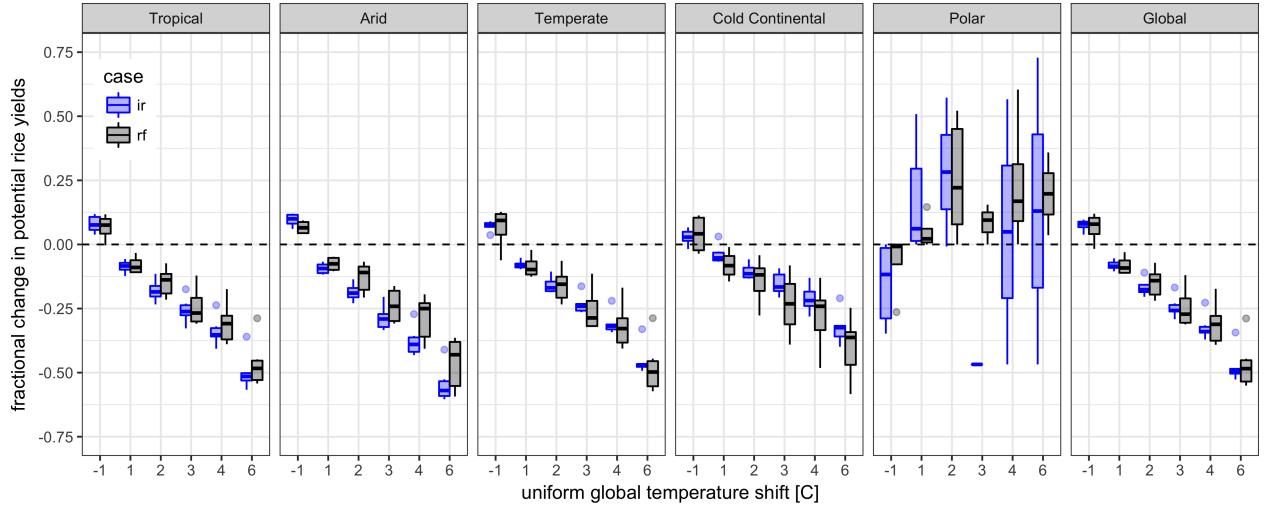
**Figure S7:** In cold continental regions, soy yields generally increase with warming and spring wheat yields decrease; the other crops are indeterminate across models. All other covariates are held constant.



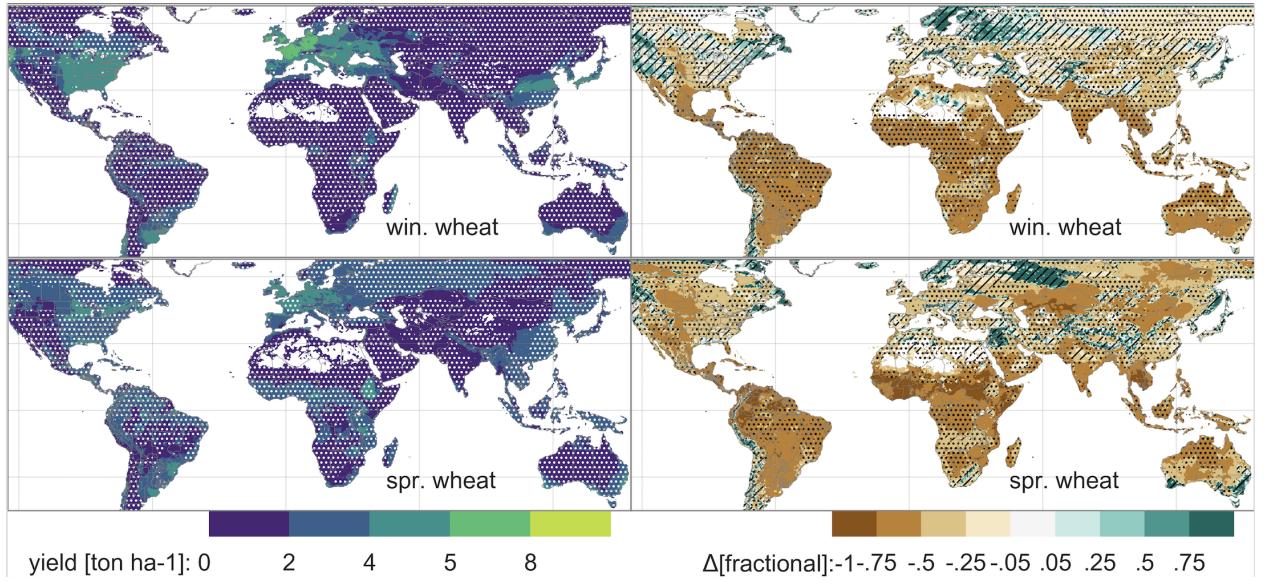
**Figure S8:** Fractional change in rice production over currently cultivated areas for rain fed and irrigated rice. All other covariates are held constant. PROMET and JULES not shown. In contrast to maize, irrigated rice is not uniformly more severely impacted at higher temperatures.



**Figure S9:** Fractional change in soy production over currently cultivated areas for rain-fed and irrigated rice. All other covariates are held constant. PROMET and JULES not shown. In contrast to maize, irrigated soy is not uniformly more severely impacted at higher temperatures.



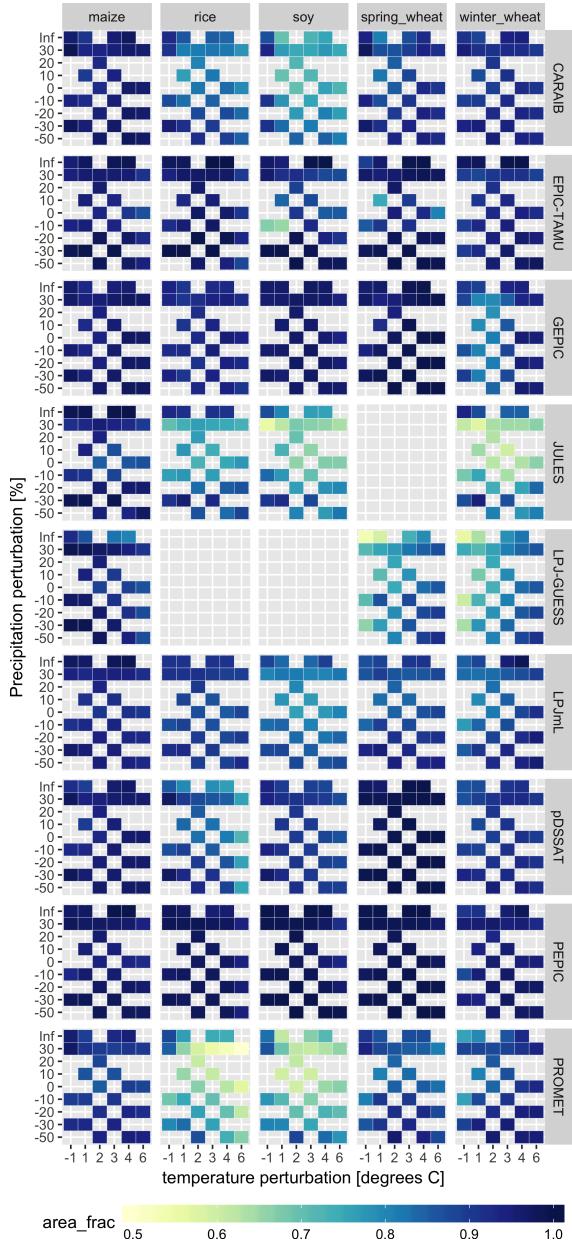
**Figure S10:** Fractional change in winter wheat production over currently cultivated areas for rain fed and irrigated rice. All other covariates are held constant. PROMET and JULES not shown.



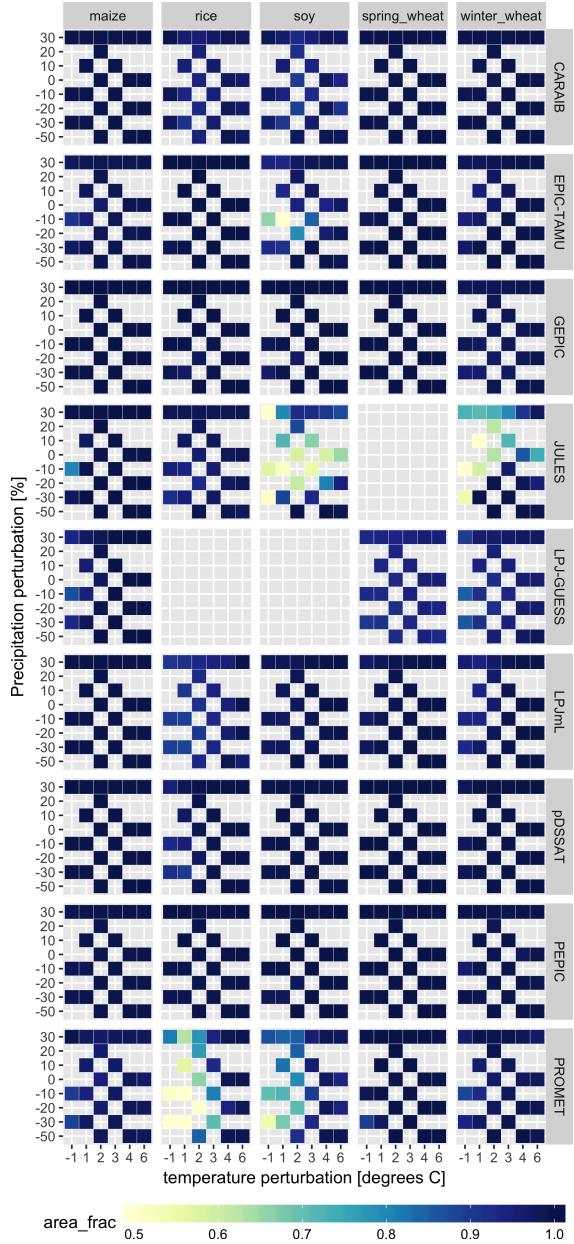
**Figure S11:** Illustration of the spatial pattern of potential yields and potential yield changes in the GGCMI Phase II ensemble, for three major crops. Left column (a) shows multi-model mean climatological yields for the baseline scenario for (top–bottom) for rain-fed wheat. White stippling indicates areas where these crops are not currently cultivated. Absence of cultivation aligns well with the lowest yield contour ( $0\text{--}2 \text{ ton ha}^{-1}$ ). Right column (b) shows the multi-model mean fractional yield change in the extreme  $T + 4 \text{ } ^\circ\text{C}$  scenario (with other inputs at baseline values). Areas without hatching or stippling are those where confidence in projections is high: the multi-model mean fractional change exceeds two standard deviations of the ensemble. ( $\Delta > 2\sigma$ ). Hatching indicates areas of low confidence ( $\Delta < 1\sigma$ ), and stippling areas of medium confidence ( $1\sigma < \Delta < 2\sigma$ ). Crop model results in cold areas, where yield impacts are on average positive, also have the highest uncertainty. Wheat is also somewhat exceptional in that also less impact in temperature and arid regions. The more complicated phenological development of winter wheat when compared to other crops is a potential source of the higher level of model disagreement.

### 3 Emulator fits and performance

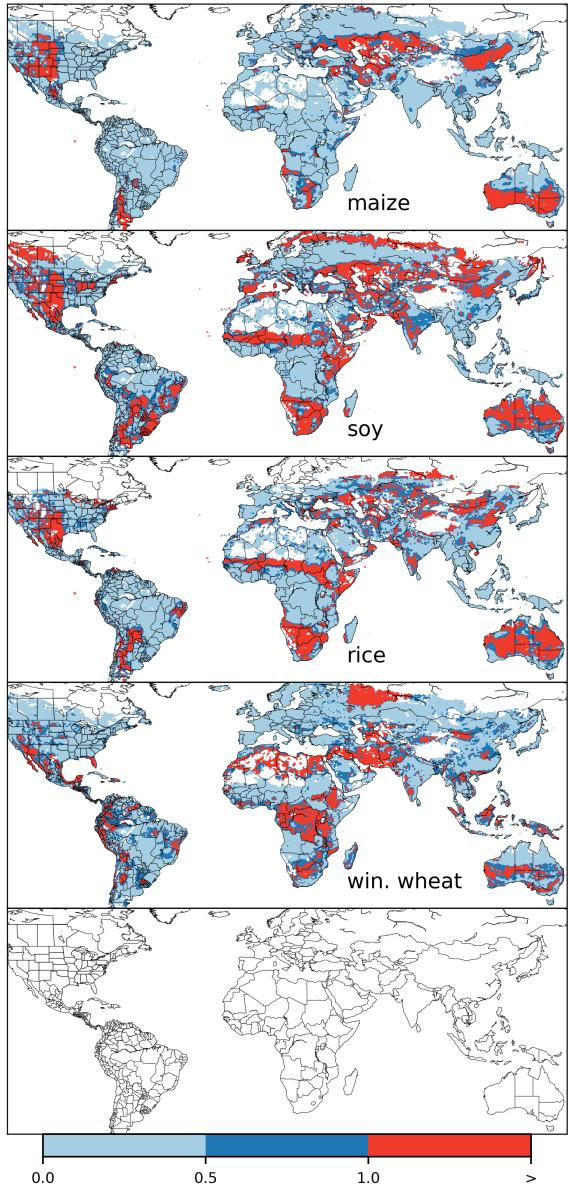
Additional Emulator performance figures for reference. PROMET and JULES shown because these two models are the most-difficult to emulate.



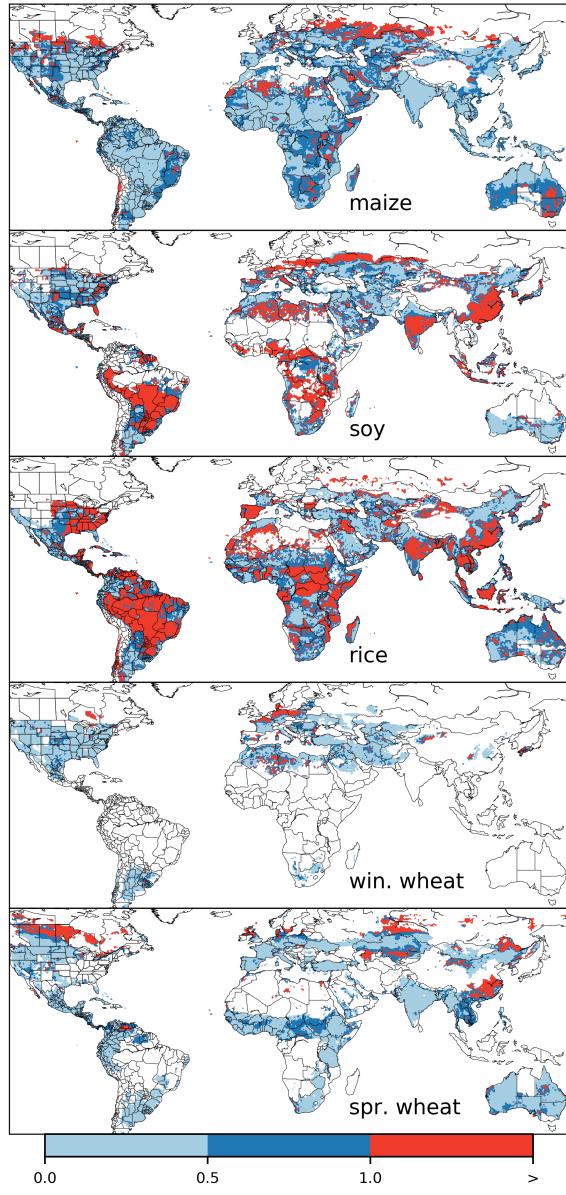
**Figure S12:** The fraction of grid cells with normalized emulation error less than 1 for the  $\text{CO}_2=360$  ppm and  $200 \text{ kg N ha}^{-1} \text{ yr}^{-1}$  case for the temperature and precipitation perturbations scenarios provided by all 9 models included in the emulator analysis. This is in contrast to the fraction of currently cultivated hectares shown in the C360 case in the manuscript and for the C810 case show in the supplemental material. The emulator is marginally more successful over currently cultivated areas than over all grid cells in general.



**Figure S13:** The fraction of currently cultivated hectares with normalized emulation error less than 1 (blue colors contours in Figure ??) for the  $\text{CO}_2=810$  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. See Equations ??, ?? for normalized error calculation. The yield response is generally easy to emulate over currently cultivated areas (dark blue and light blue).



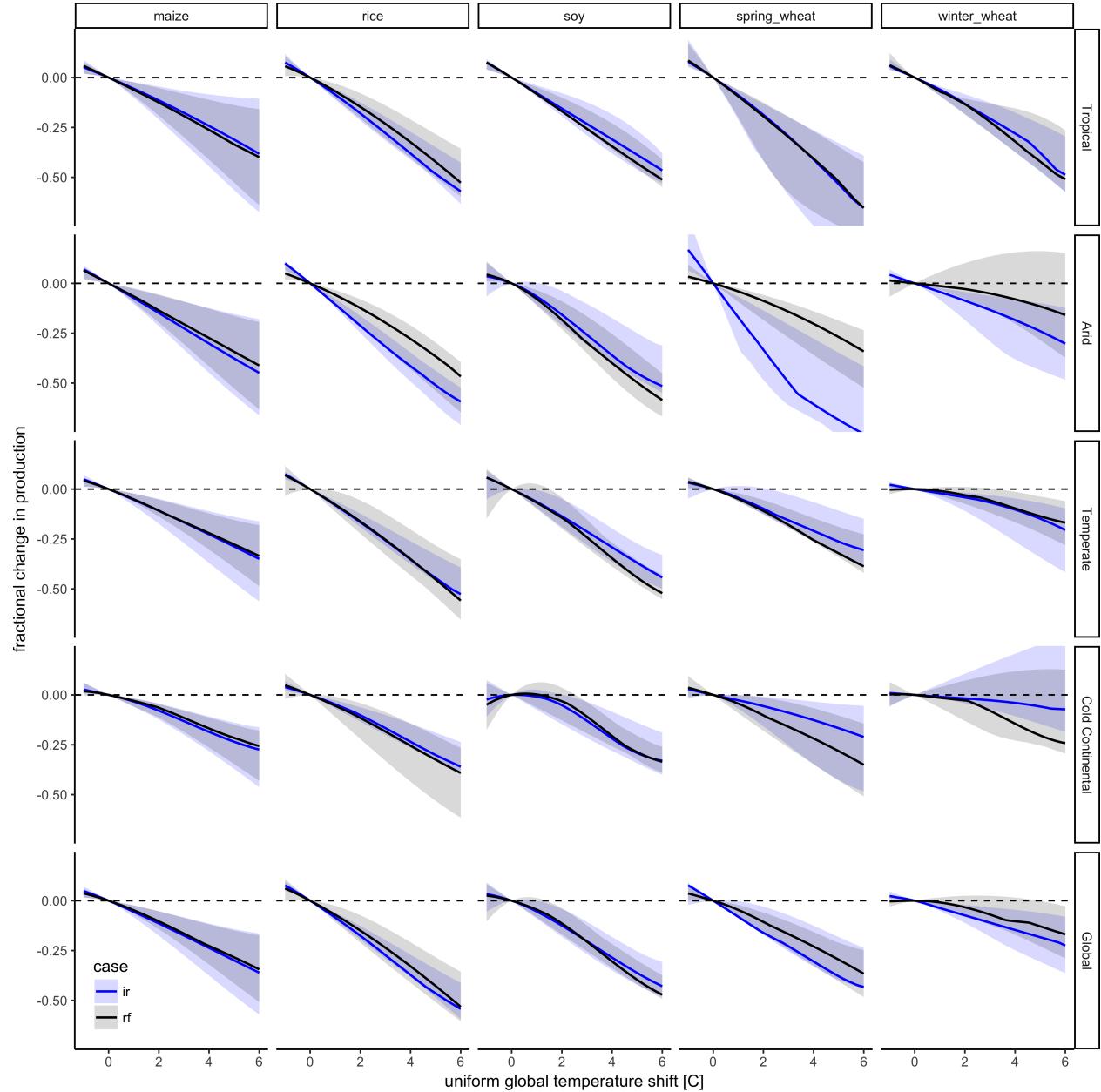
**Figure S14:** Illustration of our test of emulator performance, applied to the JULES model for the T +4 scenario for rain-fed crops. Contour colors indicate the normalized emulator error  $e$ , where  $e$  less than 1 means that emulator error exceeds the multi-model standard deviation. White areas are those where crops are not simulated by this model. Models differ in their areas omitted, meaning the number of samples used to calculate the multi-model standard deviation is not spatially consistent in all locations.



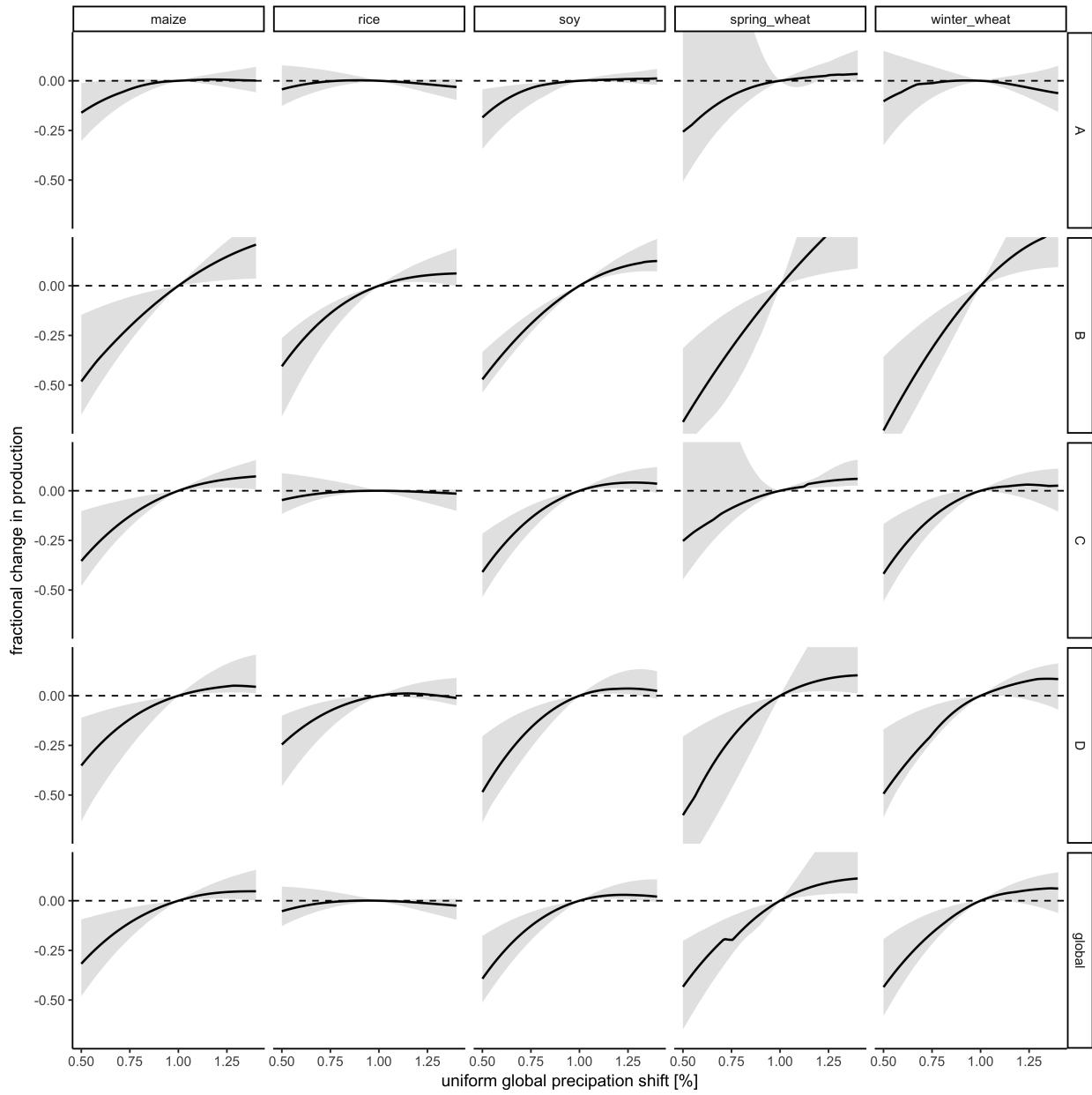
**Figure S15:** Illustration of our test of emulator performance, applied to the PROMET model for the T +4 scenario for rain-fed crops. Contour colors indicate the normalized emulator error  $e$ , where  $e$  less than 1 means that emulator error exceeds the multi-model standard deviation. White areas are those where crops are not simulated by this model. Models differ in their areas omitted, meaning the number of samples used to calculate the multi-model standard deviation is not spatially consistent in all locations.

## 4 Emulator results

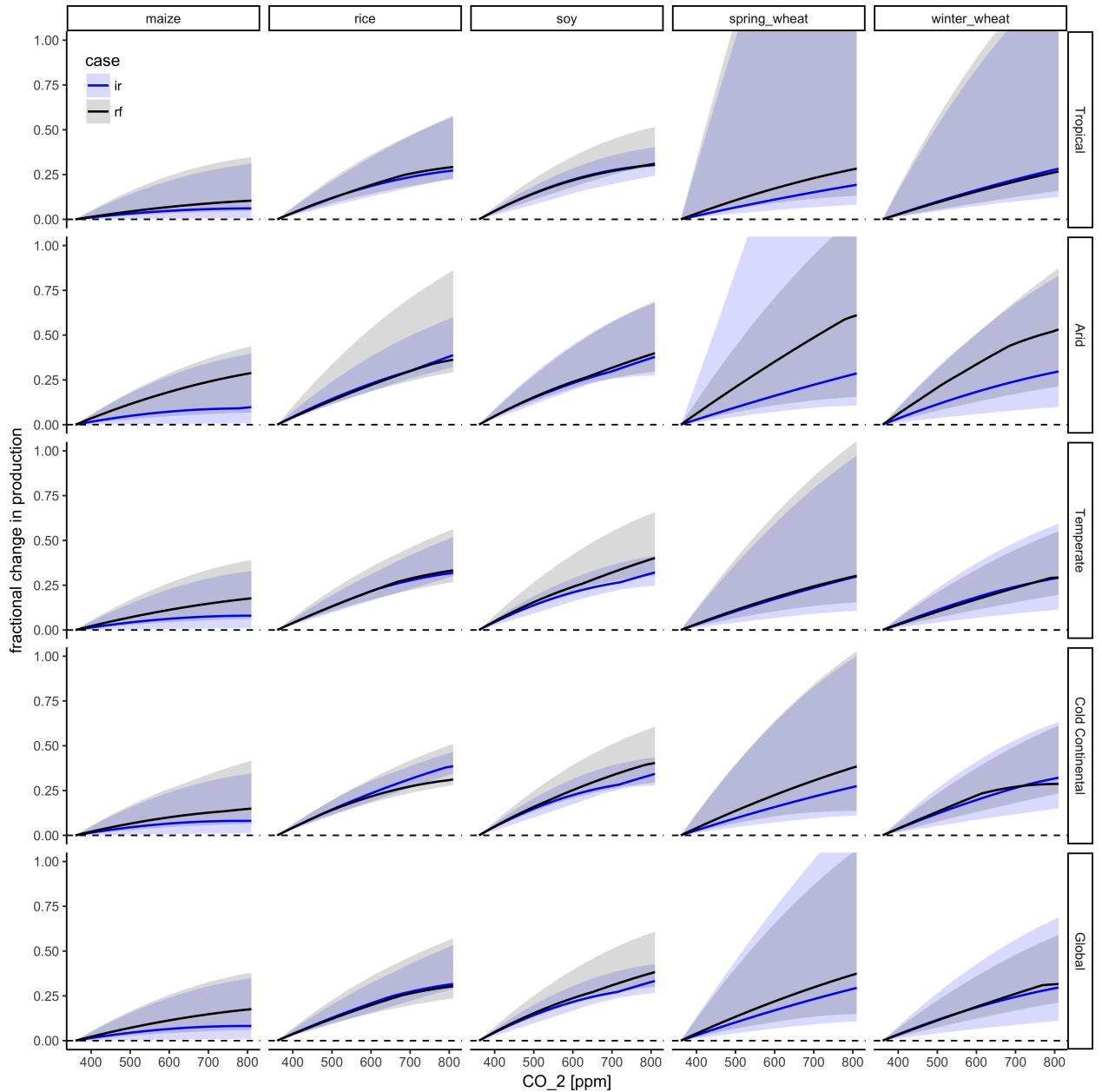
24 Example damage functions over the the four dimensions included in the study. All crops shown and split by KG climate regions.



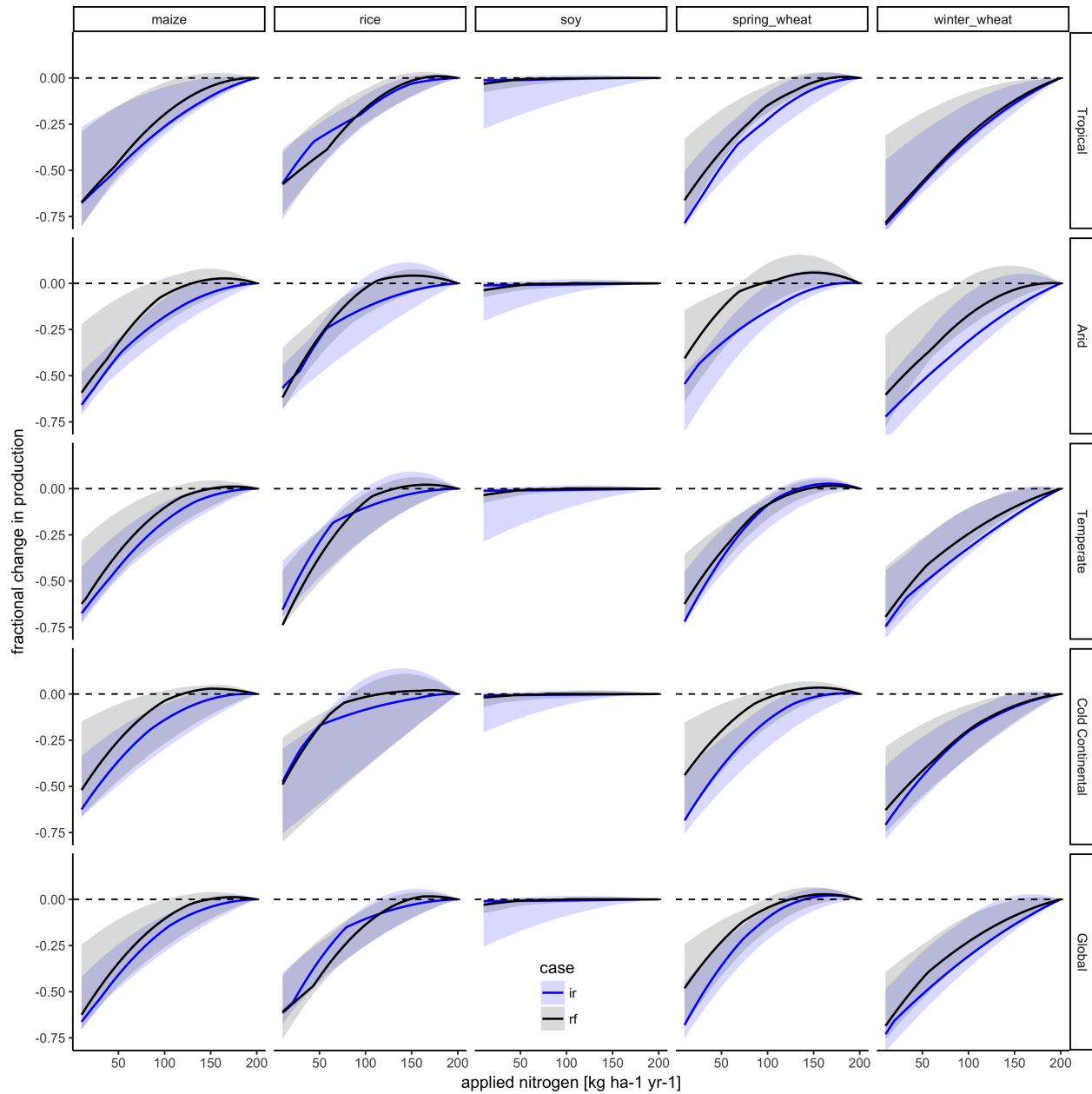
**Figure S16:** Multi-model ensemble spread in sensitivity to changes in the temperature dimension. 5% median and 95% percentile emulated damage function for currently cultivated areas. Irrigated and rain-fed crops shown. PROMET and JULES removed from ensemble. All other covariates held constant. Uniform temperature shift should not be interpreted as a realistic climate change.



**Figure S17:** Multi-model ensemble spread in sensitivity to changes in the water dimension. 5% median and 95% percentile emulated damage function for currently cultivated areas. PROMET and JULES removed from ensemble. All other covariates held constant. Uniform precipitation shift should not be interpreted as a realistic climate change.



**Figure S18:** Multi-model ensemble spread in sensitivity to changes in the carbon dimension. 5% median and 95% percentile emulated damage function for currently cultivated areas. Irrigated and rain-fed crops shown. PROMET and JULES removed from ensemble. All other covariates held constant.



**Figure S19:** Multi-model ensemble spread in sensitivity to changes in the nitrogen dimension. 5% median and 95% emulated damage function for currently cultivated areas. Irrigated and rain-fed crops shown. PROMET and JULES removed from ensemble. All other covariates held constant. Soy shows the lowest sensitivity to Nitrogen as expected. Irrigated crops are more sensitive to low Nitrogen