Strong regional influence of climatic forcing datasets on global crop model ensembles

Supplementary Material for Agricultural and Forest Meteorology

Alex C. Ruane[a], Meridel Phillips[b,a], Christoph Müller[c], Joshua Elliott[d], Jonas Jägermeyr[b,a], Almut Arneth[e], Juraj Balkovic[f,g], Delphine Deryng[h], Christian Folberth[g], Toshichika Iizumi[i], Roberto C. Izaurralde[j], Nikolay Khabarov[k], Peter Lawrence[l], Wenfeng Liu[m], Stefan Olin[n], Thomas A. M. Pugh[n,o,p], Cynthia Rosenzweig[a], Gen Sakurai[i], Erwin Schmid[q], Benjamin Sultan[r], Xuhui Wang[s,t], Allard de Wit[u], Hong Yang[v,w]

[a]NASA Goddard Institute for Space Studies, New York, New York, USA
[b]Center for Climate Systems Research, Columbia University, New York, New York, USA
[c]Potsdam Institute for Climate Impact Research, Member of the Leibniz Association, Potsdam, Germany
[d] University of Chicago, Chicago, Illinois, USA
[e] Department Atmospheric Environmental Research, Karlsruhe Institute of Technology, Garmisch-Partenkirchen, Germany
[f] Department of Soil Science, Comenius University in Bratislava, Bratislava, Slovak Republic
[g] Biodiversity and Natural Resources, International Institute for Applied Systems Analysis, Laxenburg, Austria
[h] IRI THESys, Humboldt-Universität zu Berlin, Berlin, Germany
[i] National Agriculture and Food Research Organization, Tsukuba, Ibaraki, Japan
[j] Department of Geographical Sciences, University of Maryland, College Park, MD, USA
[k] Advanced Systems Analysis Program, International Institute for Applied Systems Analysis, Laxenburg, Austria
[l] Earth System Laboratory, National Center for Atmospheric Research, Boulder, CO, USA
[m] College of Water Resources and Civil Engineering, China Agricultural University, Beijing, China
[n] Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden
[o] School of Geography, Earth & Environmental Science, University of Birmingham, Birmingham, United Kingdom
[p] Birmingham Institute of Forest Research, University of Birmingham, Birmingham, United Kingdom
[q] Institute for Sustainable Economic Development, University of Natural Resources and Life Sciences, Vienna, Austria
[r] ESPACE-DEV, Institute of Research for Development, Montpellier, France
[s] Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France
[t] Sino-French Institute of Earth System Sciences, Peking University, Beijing, China
[u] Earth Observation and Environmental Informatics, Wageningen Environmental Research, Wageningen, The Netherlands
[v] Eawag, Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland
[w] Department of Environmental Sciences, MGU, University of Basel, Basel, Switzerland

Corresponding Author's Address
Alex Ruane
2880 Broadway
New York, NY 10025, USA
alexander.c.ruane@nasa.gov

Table of Contents

S1. Phases of the Global Gridded Crop Model Intercomparison (GGCMI)

The AgMIP Global Gridded Crop Model Intercomparison (GGCMI) is a core element of the AgMIP Gridded Crop Modeling Initiative (Ag-GRID) that aims to improve and apply crop models across large spatial domains using high-performance computational resources. GGCMI community efforts have organized around multiple research phases:

- *GGCMI Fast Track* [Conducted in coordination with the Inter-Sectoral Impacts Model Intercomparison Project (ISIMIP; Warszawski et al., 2014)] – Employed 7 global gridded crop models (GGCMs) to project global yield and production changes for maize, wheat, rice, and soybean driven by 5 earth system models (ESMs) from the Fifth Coupled Model Intercomparison (CMIP5; Taylor et al., 2012) and 4 greenhouse gas emissions scenarios (Rosenzweig et al., 2014).
- *GGCMI Phase 1* [Further evaluated in this study; summarized here with further detail in the methods below (Section 2)] – Explored 1981-2010 historical period performance using 11 CFDs and 14 GGCMs simulating maize, wheat, rice, and soybean under harmonized conditions to isolate model differences (Elliott et al., 2015; Müller et al., 2019). Also used as the basis for the development of a set of GGCM historical period benchmarks to evaluate model performance and document model improvement (Folberth et al., 2019; Müller et al., 2017).
- *GGCMI Phase 2* – Sensitivity studies assessing the response of 13 GGCMs (maize, spring wheat, winter wheat, rice, and soybean) to climatological changes in carbon dioxide (CO_2), temperature, and precipitation under different levels of nitrogen fertilizer and adaptation (Franke et al., 2020, 2019; Minoli et al., 2019).
- *GGCMI Phase 3* [Anticipated 2020; in coordination with ISIMIP] – Multi-model simulations using updated GGCM versions to project climate impacts on yield and production using outputs from the Sixth Coupled Model Intercomparison Project (CMIP6; Eyring et al., 2016). These simulations are driven by scenarios based upon the W5E5 CFD (Lange, 2019).

Results from GGCMI phases build upon one another, and this study's focus on the role of climatic forcing dataset selection in capturing agricultural response to climatic variation will therefore also assist in interpretation of other GGCMI phases as well as the efficient design of future assessments.

S2. GGCM configurations

This study utilizes output from GGCMI Phase 1 experiments, which we summarize here while referring readers to the full protocols described by Elliott et al., (2015). The 14 GGCMs presented in **Table S1** simulated the 1981-2010 harvest years using the 11 CFDs described in the previous section. Maize, wheat, rice, and soybean simulations were conducted on a global, 0.5° x 0.5° grid (~55 km wide grid cells at the equator) using one of three configurations:

- *Harmonized* (H) – GGCMs use consistent fertilizer levels, irrigation guidelines, and planting dates with cultivars configured to have the same average maturity date (usually achieved by setting thermal unit requirements so that each phenological stage would complete given expected growing degree days). See Elliott et al., (2015) for further information.
- *Harmonized with no nitrogen limitation* (N) – Same as Harmonized configuration but all nitrogen response is turned off (to be consistent with models that do not resolve nitrogen stress).
- *Default* (D) – GGCMs use 'best guess' configuration for all settings used prior to introduction of harmonized settings.

Heterogeneity in cultivars, soils, and farm management (including growing season dates) below the 0.5° x 0.5° grid scale may affect comparisons with real-world crop perceptions of agro-climatic anomalies.

Benchmarks for the performance of GGCMs were established by Müller et al., (2017), who documented structural differences in the models and found significant skill in reproducing national-level production anomalies across many GGCMs when driven by the AgMERRA CFD, although particular models and regions were less skillful. Time series comparisons between CFDs and GGCMI outputs are not detrended to account for major climatic response (in Section 3) as GGCMI configurations do not change from year to year. Additional crop species simulated by a smaller subset of GGCMs within GGCMI Phase 1 are not analyzed here (Elliott et al., 2015).

S3. Additional GGCM information

The GGCMI ensemble includes models that share similar attributes even as each produces unique results. Rosenzweig et al., (2014a) noted a divergence in future climate change responses between models depending on whether they included nitrogen stress, although those differences were attributed to an interaction with the long-term rise in CO_2 concentrations that is not as prominent within the 1981-2010 period. Five GGCMs are built from the field-scale EPIC model, although Folberth et al. (2019) has documented differences in model performance and responsiveness owing to slightly different model versions, dynamic soil parameterizations, hydrology, cultivar distributions, and field management. LPJmL and LPJ-GUESS are built from the same biophysical process core but differ substantially in other process representations, while pAPSIM and pDSSAT utilize a common computational framework that aims to harmonize configuration settings and inputs to the greatest extent possible (Elliott et al., 2014).

Table S1: Overview of Global Gridded Models (GGCMs). Further details provided in Elliott et al., (2015) and Müller et al., (2019). Table reflects data accessed April 25, 2019.

Model name	Modeling Center	Notes and key reference
CGMS-WOFOST	Wageningen University, the Netherlands	Based on WOFOST site-based process model (de Wit et al., 2019; van Diepen et al., 1989; van Ittersum et al., 2003)
CLM-Crop	NCAR, USA	Dynamic vegetation model based on coupled carbon-nitrogen version of CLM (Drewniak et al., 2013)
EPIC-BOKU	Boku, Austria	Based on EPIC site-based process model (Izaurralde et al., 2006; Williams, 1995). Core EPIC model v0810
EPIC-IIASA	IIASA, Austria	Based on EPIC site-based process model, as above. Core EPIC model v0810
EPIC-TAMU	Texas A&M University, USA	Based on EPIC site-based process model, as above. Core EPIC model v1102
GEPIC	IIASA, Austria	Based on EPIC site-based process model as above (additional detail in Liu et al., 2007). Core EPIC model v0810
LPJ-GUESS	KIT, Germany; Lund Univeristy, Sweden	Crop-enabled version of LPJ-GUESS dynamic vegetation model, loosely drawing on crop parameterizations in LPJmL (Lindeskog et al., 2013; Olin et al., 2015)
LPJmL	PIK, Germany	Dynamic vegetation model (Bondeau et al., 2007; Müller and Robertson, 2014; Schaphoff et al., 2018)
pAPSIM	University of Chicago, USA	Based on APSIM site-based process model (Elliott et al., 2014; Holzworth et al., 2014; Keating et al., 2003)
pDSSAT	University of Chicago, USA	Based on DSSAT v4.5 site-based process model (Elliott et al., 2014; Jones et al., 2003)
PEGASUS	Tyndall Centre, University of East Anglia, UK	Process-based vegetation model (Deryng et al., 2014, 2011)
ORCHIDEE-CROP	IPSL, France	Includes model processes based on STICS site-based process model (Wu et al., 2016)

PEPIC	Eawag, Switzerland	Based on EPIC site-based process model (Liu et al., 2016a, 2016b; Williams et al., 1984)
PRYSBI2	NAIES, Japan	Empirical/process-based hybrid model (Okada et al., 2015; Sakurai et al., 2014)

S4. Production datasets and processing

GGCMs simulate crop yields (t/ha) that must be converted to production (total kg) using harvested area masks in order to compare against observational production datasets. We calculate national-level production from the 0.5° x 0.5° grid using harvested crop areas from the Spatial Production Allocation Model v2.0 (SPAM), which approximates the year 2005 and does not change from year to year (You et al., 2014). We aggregate rainfed and irrigated production values separately using the corresponding GGCMI simulations and SPAM areas, then use the sum of rainfed and irrigated production for national or global totals (following Ruane et al., 2018b). Porwollik et al., (2017) showed that choice of harvested area dataset used to weight different GGCMI grid cells in production calculations matters for many GGCM/crop/country combinations. Resolving these differences is beyond the purview of this study, but it is likely that improved area and growing season data would improve our analyses. Due to inconsistencies between individual crop models' reported harvest years and FAO reporting methods, we shift the time series by one year in either direction and use correlations for the shifted time series if the correlation improves by at least 0.2 (a similar approach was followed by Müller et al. 2017). Iizumi et al. (2017a) alternatively used a 2-year running average to address this inconsistent harvest year reporting problem.

S5. GGCMI historical production and top producing countries

Figure S.1 shows major production regions in the world according to GGCMI rainfed simulations and FAO statistics. Areas with the highest yields do not always correspond to the highest planted areas, however; due to the availability of irrigation, the overall amount of arable land, and competition from other potential land uses. A discussion of mean production biases across GGCMI Phase 1 is beyond the purview of this study and was the subject of in-depth analysis by Müller et al. (2017). Global production is highly dependent on relatively few major production regions for each crop, although communities with low yields or production can still be highly dependent on reliable harvests in order to maintain food security and agricultural sector revenues crucial to economic vibrancy. Model evaluation focuses on the top 20 national producers for each crop commodity according to the 2013-2017 period.

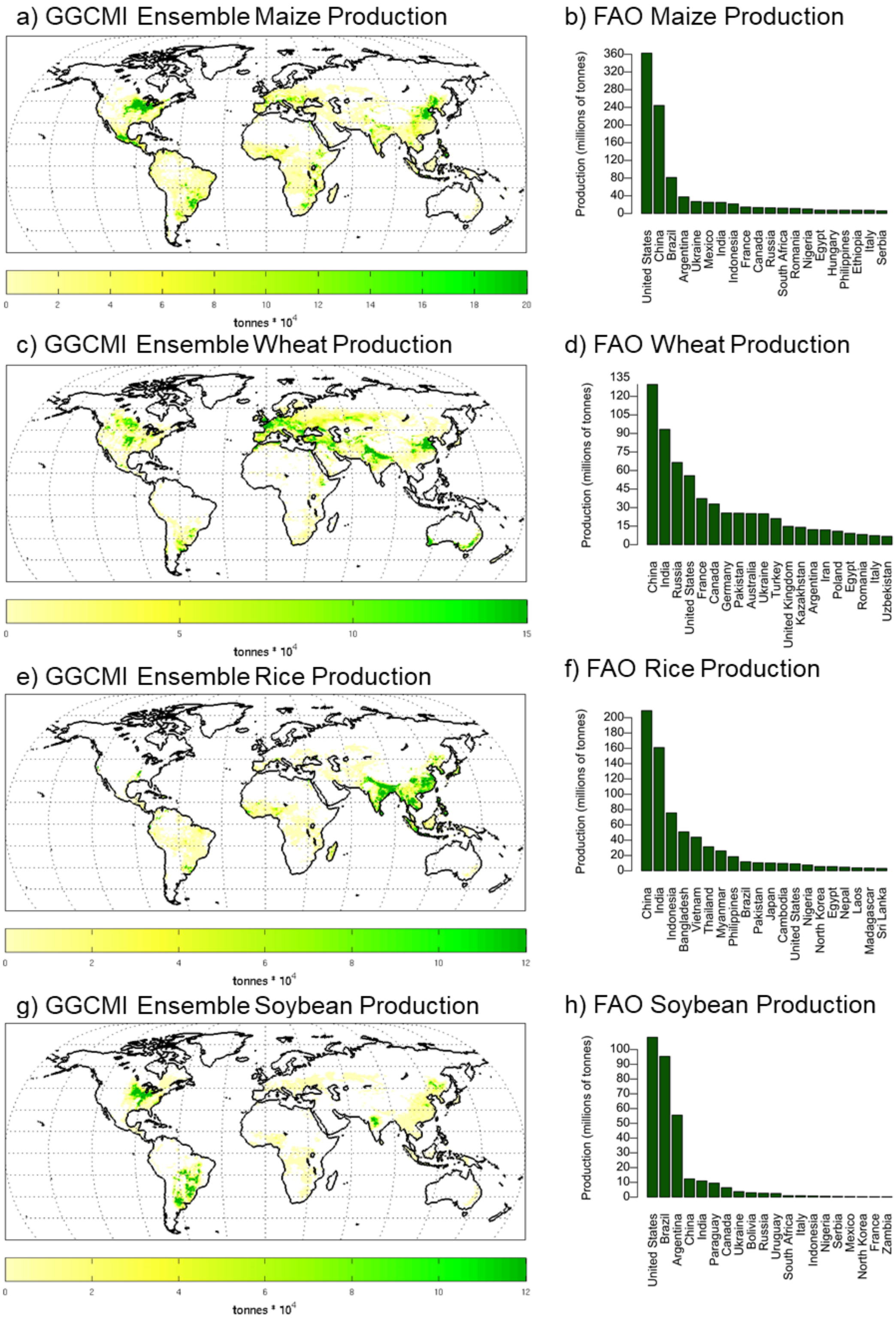

Figure S.1: GGCMI *Ensemble-all* yield (1981-2010) and top 20 FAO national producers (2013-2017) for (a,b) maize; (c,d) wheat; (e,f) rice; and (g,h) soybean. Areas with fewer than 10 ha harvested area for the given rainfed crop are omitted from maps in order to highlight major agricultural regions (You et al., 2014).

Figure S.2 provides an example of GGCM/CFD ensembles evaluated in this study. These include an example of a single GGCM/CFD combination (EPIC-TAMU using the AgMERRA CFD), additional models included in *AgMERRA+* (7 GGCMs total), additional models in *AgMERRA-all* (14 GGCMs total), and additional models in the *Ensemble-All* "Super Ensemble" (all 91 GGCM/CFD combinations), as well as the ensemble mean of each ensemble. Similar to results in other countries and crop species, these results show the improvement of correlations in crop model ensembles (**Figures 4c, 6a**), with large improvement in the creation of a multi-GGCM ensemble compared to the larger range of individual models even as the creation of larger ensembles has a diminishing benefit.

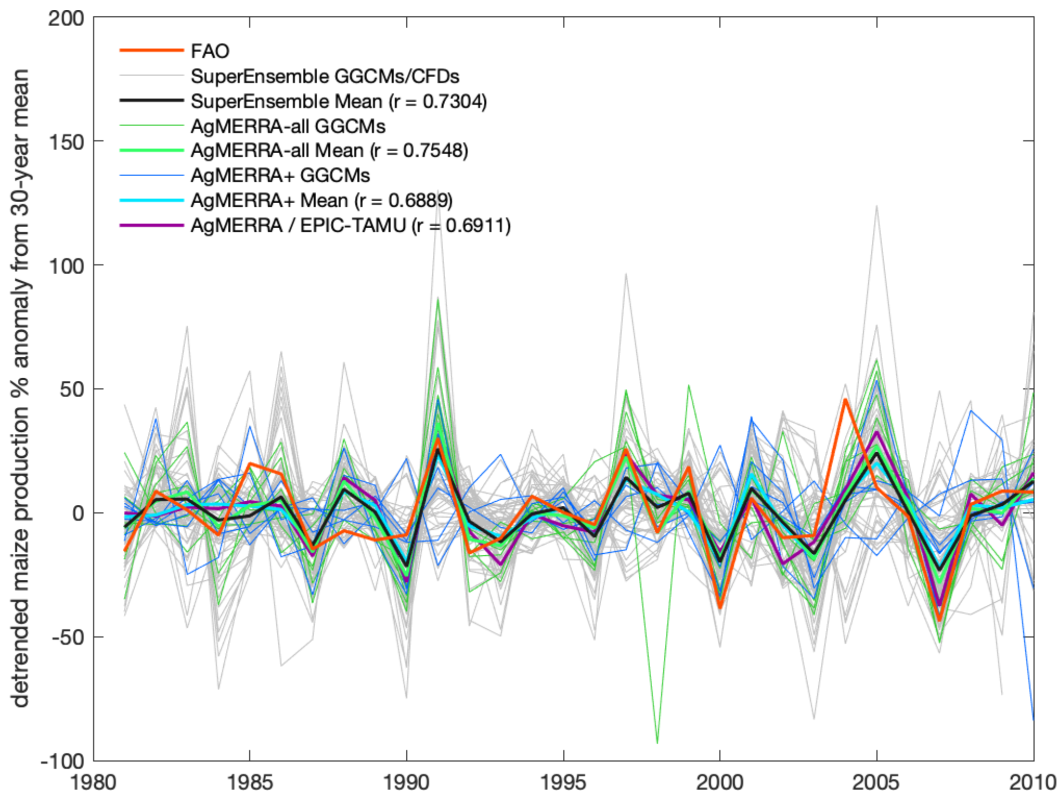

Figure S.2: Example of ensemble construction for Romanian maize production anomalies (compared to mean of 1981-2010), including detrended FAO observations and various GGCM/CFD simulation ensemble sets evaluated in this study.

S6. Relative bias of climatic forcing datasets

Figure 1 shows the *CFD-all* ensemble mean rainfed maize season (left column) and relative deviations for the two most commonly simulated CFDs (AgMERRA and WFDEIgpcc) for mean temperature, precipitation, and solar radiation, as well as extreme heat, rainy day, and heavy rainfall day metrics. **Figures S.3-S.5** show corresponding bias maps (relative to *CFD-all*) for the other CFDs analyzed within this study. We include the W5E5 climatic forcing dataset in Figure S.5 as this CFD is the basis of GGCMI Phase 3 simulations.

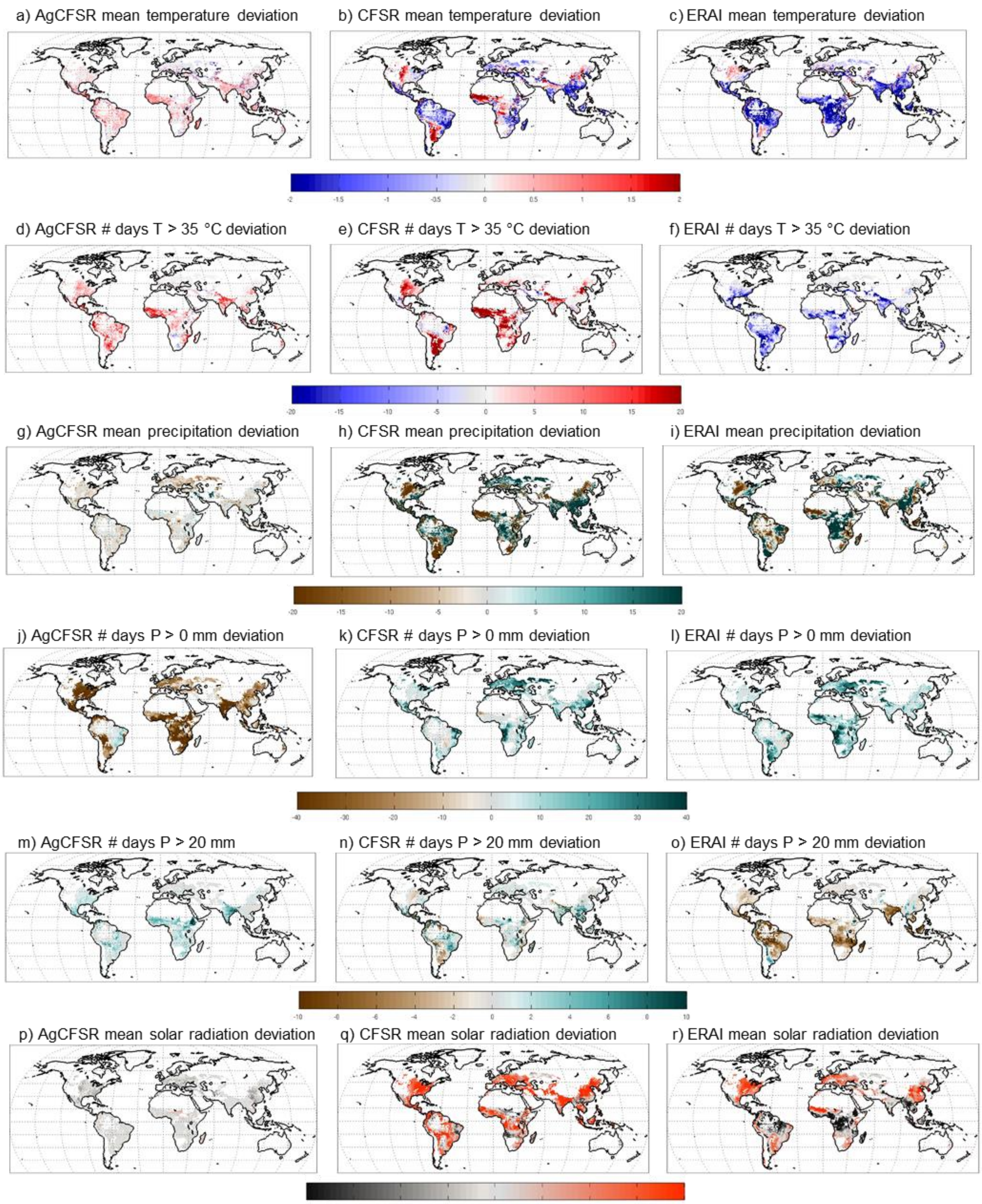

Figure S.3: Extension of Figure 1 for additional climatic forcing datasets (CFDs). Rainfed maize growing season deviation (1981-2010) compared to *CFD-all* mean for (left) AgCFSR, (center) CFSR, and (right) ERAI. From top to bottom, rows are deviations in growing season mean temperature (°C), mean precipitation (%), mean solar radiation (MJ m^{-2} day^{-1}), mean number of days where Tmax > 35 °C, mean number of days where P > 0 mm/day, mean number of days where P > 20 mm/day.

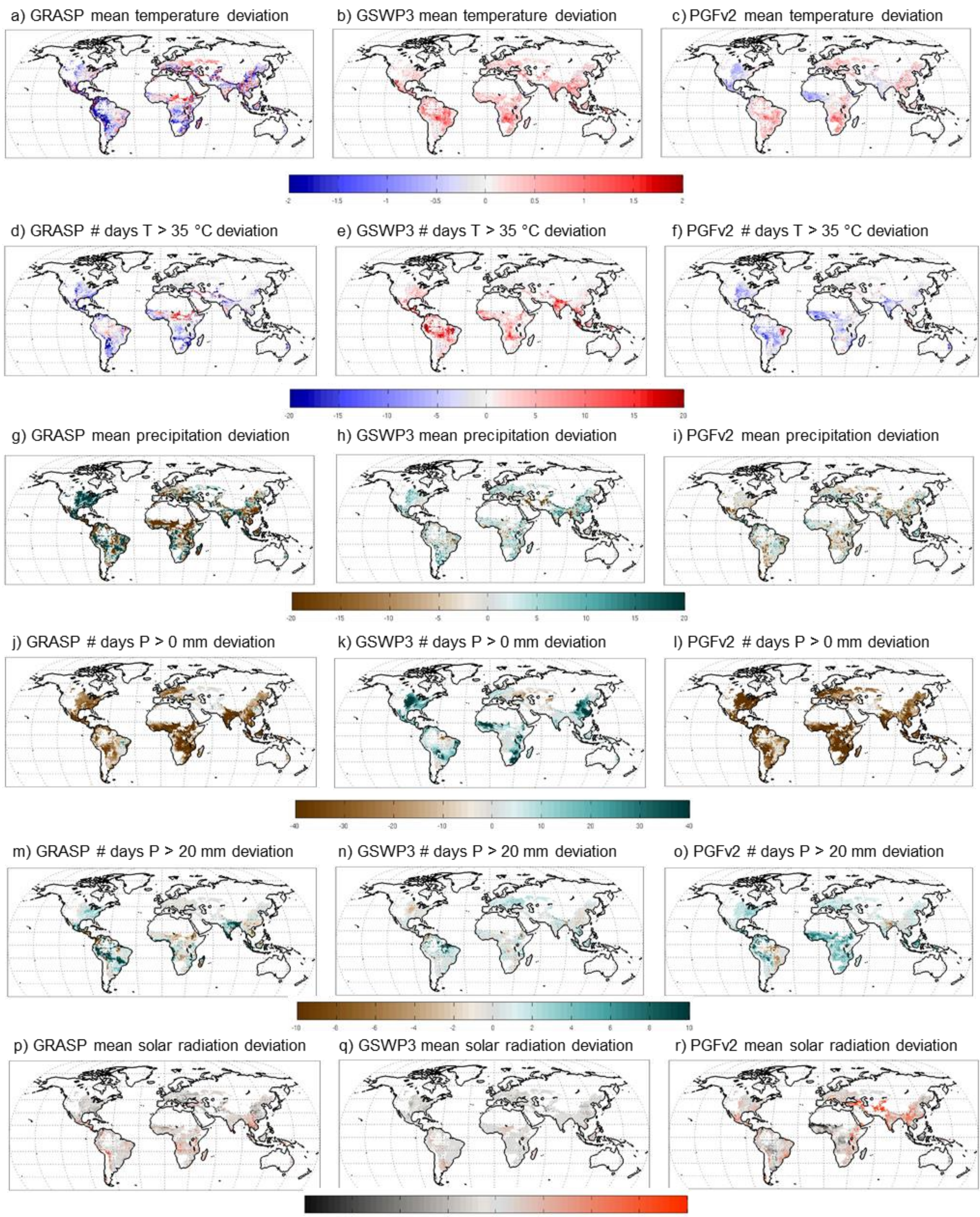

Figure S.4: Extension of Figure 1 for additional climatic forcing datasets (CFDs). Rainfed maize growing season deviation (1981-2010) compared to *CFD-all* mean for (left) GRASP, (center) GSWP3, and (right) PGFv2. From top to bottom, rows are deviations in growing season mean temperature (℃), mean precipitation (%), mean solar radiation (MJ m^{-2} day^{-1}), mean number of days where Tmax > 35 °C, mean number of days where P > 0 mm/day, mean number of days where P > 20 mm/day.

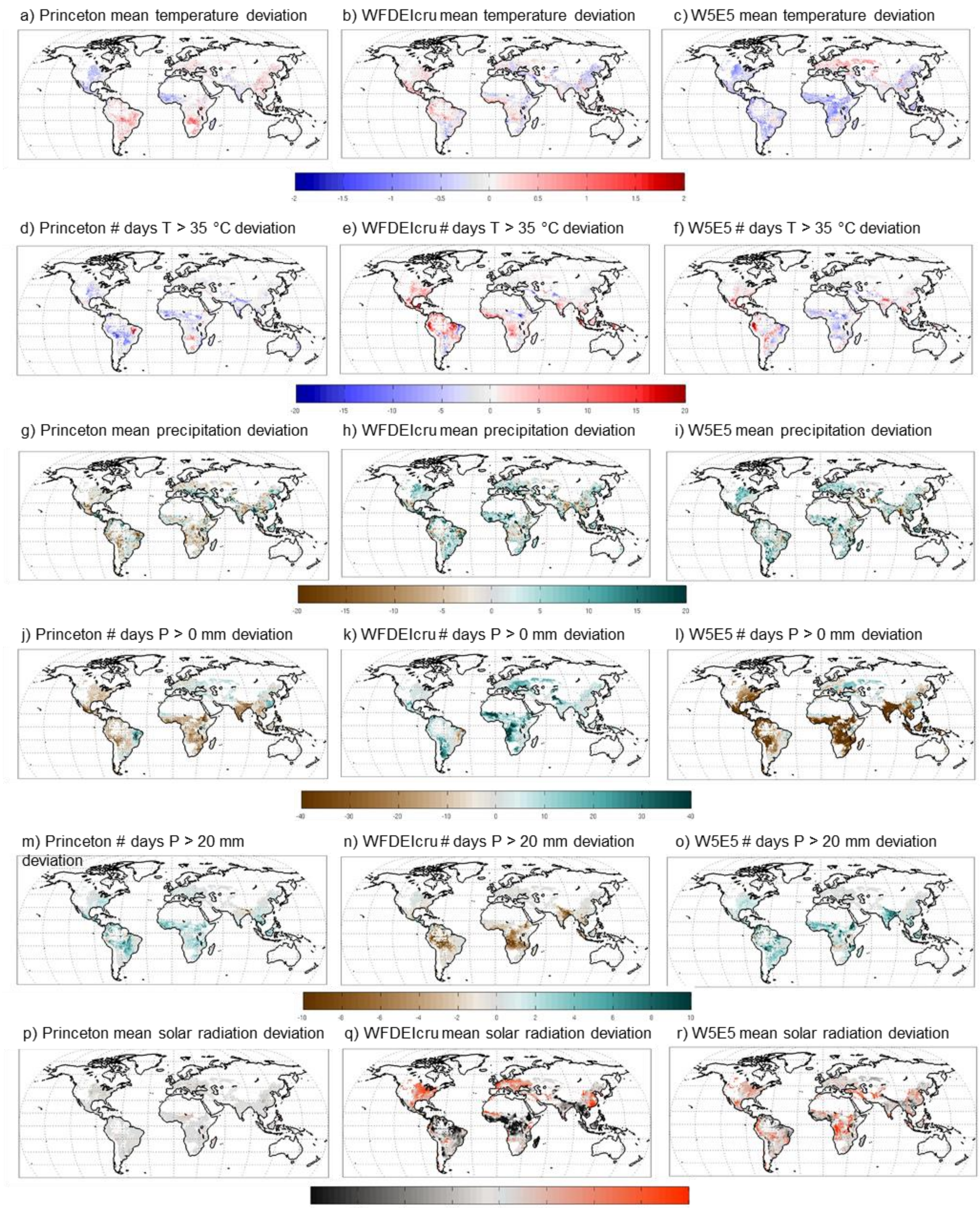

Figure S.5: Extension of Figure 1 for additional climatic forcing datasets (CFDs). Rainfed maize growing season deviation (1981-2010) compared to *CFD-all* mean for Princeton (left), WFDEIcru (center) and W5E5 (right). From top to bottom, rows are deviations in growing season mean temperature (℃), mean precipitation (%), mean solar radiation (MJ m^{-2} day^{-1}), mean number of days where Tmax > 35 °C, mean number of days where P > 0 mm/day, mean number of days where P > 20 mm/day. Checkerboard pattern in Princeton # days above P > 0 mm/day appears to be a result of re-gridding to GGCMI 0.5˚ x 0.5˚ resolution.

S7. Standardized Anomalies

Figure S.6 shows two examples of the standardized anomaly calculation using rainfed soybean precipitation for sites in Nigeria and the United States. The standardized anomaly compares uncertainty across the CFDs to uncertainty of the true signal (assumed here to be the CFD ensemble mean) across the years. The former is illustrated by showing each CFD's precipitation (gray lines), the CFD ensemble standard deviation each year (red whiskers), and the overall standard deviation of CFD annual anomalies across all years (red triangle in blue box on right). The latter is illustrated by presenting the ensemble average CFD anomaly time series (black line) and its overall standard deviation across all years (black triangles in blue box on right). As indicated in the main text, standardized precipitation anomalies are the ratio of (i) the standard deviation of yearly CFD anomalies (compared to the CFD ensemble mean) to (ii) the standard deviation of the CFD ensemble mean time series itself, represented as (i) the red triangle divided by (ii) the black triangle within the blue box on the right side of each Figure S.6 panel. Taraba, Nigeria, has a high standardized precipitation anomaly due to large variation among CFDs and a small interannual variation in the CFD ensemble average (typical of many tropical climates with sparse observational networks). Adams County, Illinois, USA, has a small standardized precipitation anomaly as the difference between CFDs is small compared to the large interannual variability in the CFD ensemble (typical of many mid-latitude continental climates with dense and high-quality observational networks).

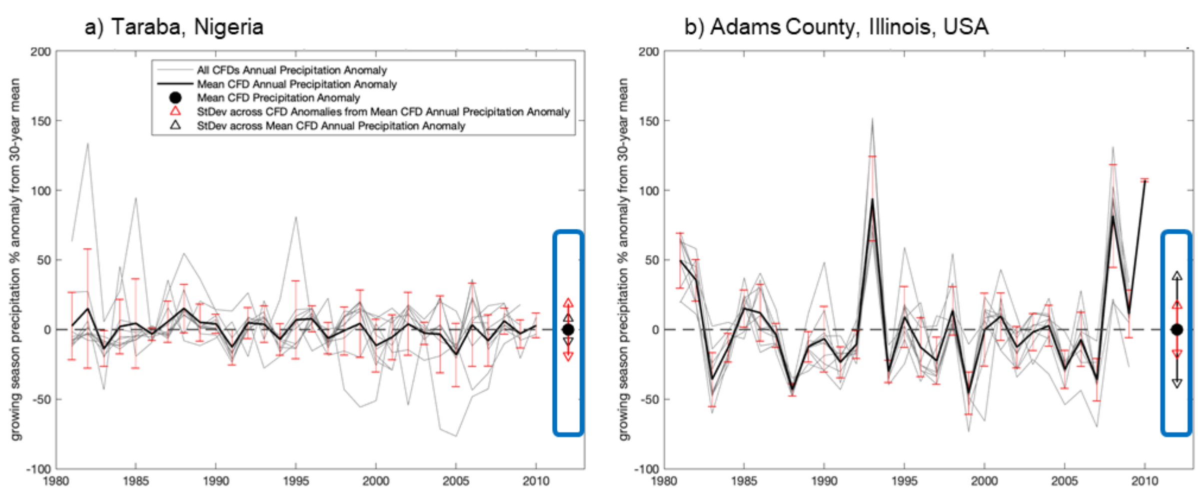

Figure S.6: Examples of standardized anomaly calculation for rainfed soybean precipitation in grid cell corresponding to a) Taraba, Nigeria, which has a high standardized anomaly, and (b) Adams County, Illinois, USA, which has a low standardized anomaly. These examples correspond to values presented in Figure S.7d below. Using the graphical representation here, the standardized anomaly is calculated as the ratio of the red triangle to the black triangle within the blue box.

To complete the analysis of standardized anomalies shown for rainfed maize and rainfed rice in Figure 2, Figure S.7 provides standardized anomalies for rainfed wheat and rainfed soybean. Most major wheat areas have low CFD uncertainty compared with the major growing areas of maize, soybean, and rice. High standardized anomalies for wheat yield in portions of Northeastern Europe indicate substantial crop model uncertainties.

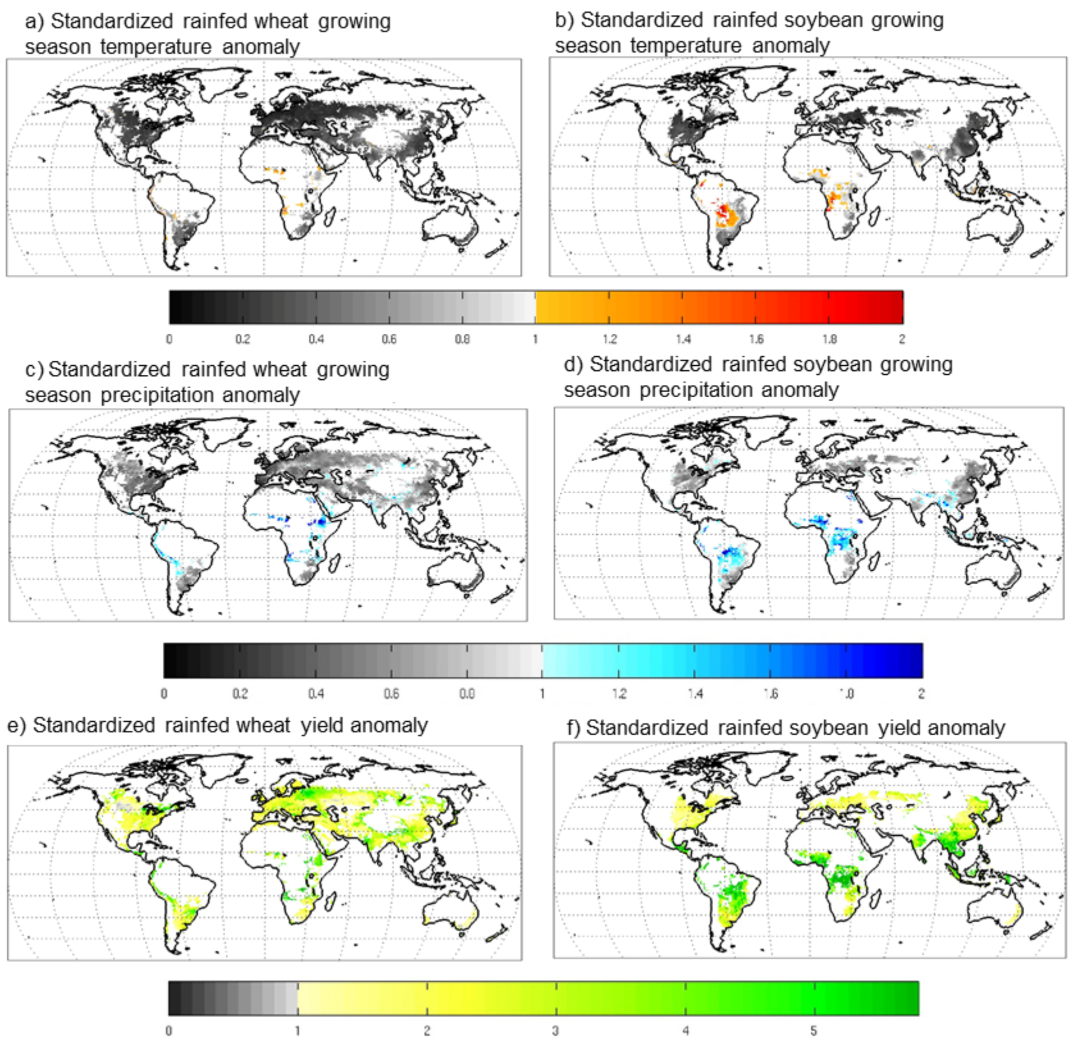

Figure S.7: Standardized anomalies (unitless) for 1981-2010 rainfed wheat growing season (left) and rainfed soybean growing season (right) mean (a,b) temperature and (c,d) precipitation (across all climatic forcing datasets) as well as for (d,e) yield (across all GGCMI x CFD combinations). Standardized anomalies are the ratio of (i) the standard deviation of yearly ensemble member anomalies (compared to the ensemble mean) to (ii) the standard deviation of the ensemble mean time series itself.

S8. Key sources of uncertainty

Uncertainty in the global gridded crop models, FAO data, and analysis methods likely obscures additional information that could be pulled from the GGCMI historical period intercomparison:

- *Model and observational coarseness* – A lack of agricultural observations restricts GGCM configurations to more regional cropping system which may differ in their responsiveness to climate. Likewise, regional production observations are not publicly released in many countries, which forced a reliance on national FAO data that aggregates across heterogeneous farming systems and unique regional environments (thus muddling the climate signal; Ray et al., 2015). GGCMs are capable of much finer resolution simulation where information is available, and perform well in places like the United States where county-level data are provided (Figure 5; Elliott et al., 2018).
- *Model harmonization* – GGCMI configurations were harmonized for fertilizer applications and growing periods, but additional model differences such as information on soil type likely interact with CFD characteristics and confound a clean climate intercomparison (this effect is reduced through the use of GGCM ensembles, *CFD+*, *CFD-all*, *Ensemble+*, and *Ensemble-all*).
- *Detrending FAO production* – Detrending FAO data against a 5-year moving average produced interannual anomalies relative to a snapshot of farm conditions in recognition of shifts in farm technologies and socioeconomic conditions. In some occasions this causes spurious results as the narrow window may be skewed by a particularly large anomaly which biases the 5-year mean. For example, years surrounding a major drought (and low production year) will be compared to a lower overall 5-yr mean production baseline that can make average years appear as 'good' years. We assume that this effect is small in comparison to the overall interannual variance captured by this detrending technique even as it is possible that some countries and years are more substantially affected. Here we utilize the same moving-window detrending approach for all countries to ensure consistency; however, development of regionally-specific detrending methods may be useful for future studies given that countries can feature strong and diverse patterns in the shape of long-term trends. The utility of this detrending approach is also evident in the high Figure 6 correlations in countries where FAO statistics had large background trends that were removed in the detrending (e.g., maize in Mexico, rice in Philippines, soybean in Paraguay).
- *Correlation* – The use of correlation as a metric for capturing climate variation provides only a partial perspective on agroclimatic limitation, with an implicit assumption that both positive and negative anomalies have a mirrored response on yields. Stakeholders may seek the hit rate statistic to anticipate the worst years, as some climate characteristics are important under particular conditions (e.g., Glotter et al. (2016) found that the number of precipitation days helps predict maize yields only when total precipitation totals were low).

S9. Opportunities for production simulation improvement and further applications

Improvement in the observed yield variability signal captured by crop models comes from four primary areas for improvement.

- *Model processes* – The inclusion of additional physical, chemical and biophysical processes that allow crop models to respond to mean climate and climate extremes. Priority development for climate-responsive GGCMs include improved representation of water logging, flood inundation (Li et al., 2019), frost and cold season damages, temperature extreme interactions with key crop development stages, and climate-driven outbreaks of pests and diseases, and overall farm management practices (e.g., Jones et al., 2017; Wang et al., 2017) as well as improved representation of phenological and allocation responses that determine yield variations (Zhu et al., 2019).

- *Model parameters* – The calibration of parameterizations that capture important responses and processes that are not directly simulated by the GGCMs (e.g., Kimball et al., 2019).
- *Configuration data* – Model settings designed to represent the local system, including growing season calendars (e.g., planting and harvesting), soil characteristics (e.g., texture), cultivar genetic traits (e.g., phenological growing degree-day requirements, radiation use efficiency), field management (e.g., tillage, planting density, planting dates), multi-cropping, and field amendments (e.g., fertilizer application, pesticides, irrigation) (e.g., (Jägermeyr and Frieler, 2018). This is particularly important for developing countries with heterogeneous small-holder farming areas.
- *Driving data* – Time series information about the field environment including the climatic forcing dataset (this study) as well as initial soil water, carbon and nitrogen pools (Basso et al., 2018). Better representation of heavy precipitation extremes (beyond 20 mm/day) will likely require higher resolution models, as the spatial averaging within these CFDs' 0.5 x 0.5 grid cells reduces high tail events.
- *Food security* – GGCMI analyses also facilitate analysis beyond the top 20 national producers, which includes many countries with low production despite a high reliance on agricultural systems for local food security and socioeconomic wellbeing.

References

Basso, B., Dumont, B., Maestrini, B., Shcherbak, I., Robertson, G.P., Porter, J.R., Smith, P., Paustian, K., Grace, P.R., Asseng, S., Bassu, S., Biernath, C., Boote, K.J., Cammarano, D., De Sanctis, G., Durand, J.-L., Ewert, F., Gayler, S., Hyndman, D.W., Kent, J., Martre, P., Nendel, C., Priesack, E., Ripoche, D., Ruane, A.C., Sharp, J., Thorburn, P.J., Hatfield, J.L., Jones, J.W., Rosenzweig, C., 2018. Soil Organic Carbon and Nitrogen Feedbacks on Crop Yields under Climate Change. Agric. Environ. Lett. 3, 180026. https://doi.org/10.2134/ael2018.05.0026

Bondeau, A., Smith, P.C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-campen, H., Müller, C., Reichstein, M., Smith, B., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Glob. Chang. Biol. 13, 679–706. https://doi.org/10.1111/j.1365-2486.2006.01305.x

de Wit, A., Boogaard, H., Fumagalli, D., Janssen, S., Knapen, R., van Kraalingen, D., Supit, I., van der Wijngaart, R., van Diepen, K., 2019. 25 years of the WOFOST cropping systems model. Agric. Syst. https://doi.org/10.1016/j.agsy.2018.06.018

Deryng, D., Conway, D., Ramankutty, N., Price, J., Warren, R., 2014. Global crop yield response to extreme heat stress under multiple climate change futures. Environ. Res. Lett. 9, 034011. https://doi.org/10.1088/1748-9326/9/3/034011

Deryng, D., Sacks, W.J., Barford, C.C., Ramankutty, N., 2011. Simulating the effects of climate and agricultural management practices on global crop yield. Global Biogeochem. Cycles 25. https://doi.org/10.1029/2009gb003765

Drewniak, B., Song, J., Prell, J., Kotamarthi, V.R., Jacob, R., 2013. Modeling agriculture in the Community Land Model. Geosci. Model Dev. 6, 495–515. https://doi.org/10.5194/gmd-6-495-2013

Elliott, J., Glotter, M., Ruane, A.C., Boote, K.J., Hatfield, J.L., Jones, J.W., Rosenzweig, C., Smith, L.A., Foster, I., 2018. Characterizing agricultural impacts of recent large-scale US droughts and changing technology and management. Agric. Syst. 159, 275–281. https://doi.org/10.1016/J.AGSY.2017.07.012

Elliott, J., Kelly, D., Chryssanthacopoulos, J., Glotter, M., Jhunjhnuwala, K., Best, N., Wilde, M., Foster, I., 2014. The parallel system for integrating impact models and sectors (pSIMS). Environ. Model. Softw. 62, 509–516. https://doi.org/10.1016/j.envsoft.2014.04.008

Elliott, J., Müller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K.J., Büchner, M., Foster, I., Glotter, M., Heinke, J., Iizumi, T., Izaurralde, R.C., Mueller, N.D., Ray, D.K., Rosenzweig, C., Ruane, A.C., Sheffield, J., 2015. The Global Gridded Crop Model Intercomparison: data and modeling protocols for Phase 1 (v1.0), Geoscientific Model Development. Copernicus GmbH. https://doi.org/10.5194/gmd-8-261-2015

Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E., 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci. Model Dev. 9, 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Folberth, C., Elliott, J., Müller, C., Balkovič, J., Chryssanthacopoulos, J., Izaurralde, R.C., Jones, C.D., Khabarov, N., Liu, W., Reddy, A., Schmid, E., Skalský, R., Yang, H., Arneth, A., Ciais, P., Deryng, D., Lawrence, P.J., Olin, S., Pugh, T.A.M., Ruane, A.C., Wang, X., 2019. Parameterization-induced uncertainties and impacts of crop management harmonization in a global gridded crop model ensemble, PLoS ONE. https://doi.org/10.1371/journal.pone.0221862

Franke, J., Müller, C., Elliott, J., Ruane, A., Jagermeyr, J., Balkovic, J., Ciais, P., Dury, M., Falloon, P., Folberth, C., Francois, L., Hank, T., Hoffmann, M., Jacquemin, I., Jones, C., Khabarov, N., Koch, M., Li, M., Liu, W., Olin, S., Phillips, M., Pugh, T.A., Reddy, A., Wang, X., Williams, K., Zabel, F., Moyer, E., 2019. The GGCMI Phase II experiment: global gridded crop model simulations under uniform changes in CO2, temperature, water, and nitrogen levels (protocol version 1.0). Geosci. Model Dev. Discuss. 1–30. https://doi.org/10.5194/gmd-2019-237

Franke, J., Müller, C., Elliott, J., Ruane, A.C., Jagermeyr, J., Snyder, A., Dury, M., Falloon, P., Folberth, C., Francois, L., Hank, T., Izaurralde, R.C., Jacquemin, I., Jones, C., Li, M., Liu, W., Olin, S., Phillips, M.M., Pugh, T.A.M., Reddy, A.D., Williams, K., Wang, Z., Zabel, F., Moyer, E.J., 2020. The GGCMI phase II emulators : global gridded crop model responses to changes in CO2, temperature, water, and nitrogen. Geosci. Model Dev. Discuss. https://doi.org/10.5194/gmd-2019-365

Glotter, M.J., Moyer, E.J., Ruane, A.C., Elliott, J., 2016. Evaluating the sensitivity of agricultural model performance to different climate inputs. J. Appl. Meteorol. Climatol. 55, 579–594. https://doi.org/10.1175/JAMC-D-15-0120.1

Holzworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D., Brown, H., Whish, J.P.M., Verrall, S., Fainges, J., Bell, L.W., Peake, A.S., Poulton, P.L., Hochman, Z., Thorburn, P.J., Gaydon, D.S., Dalgliesh, N.P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F.Y., Wang, E., Hammer, G.L., Robertson, M.J., Dimes, J.P., Whitbread, A.M., Hunt, J., van Rees, H., McClelland, T., Carberry, P.S., Hargreaves, J.N.G., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., Keating, B.A., 2014. APSIM - Evolution towards a new generation of agricultural systems simulation. Environ. Model. Softw. 62, 327–350. https://doi.org/10.1016/j.envsoft.2014.07.009

Iizumi, T., Furuya, J., Shen, Z., Kim, W., Okada, M., Fujimori, S., Hasegawa, T., Nishimori, M., 2017. Responses of crop yield growth to global temperature and socioeconomic changes. Sci. Rep. 7, 7800. https://doi.org/10.1038/s41598-017-08214-4

Izaurralde, R.C., Williams, J.R., McGill, W.B., Rosenberg, N.J., Jakas, M.C.Q., 2006. Simulating soil C dynamics with EPIC: Model description and testing against long-term data. Ecol. Modell. 192, 362–384. https://doi.org/10.1016/j.ecolmodel.2005.07.010

Jägermeyr, J., Frieler, K., 2018. Spatial variations in crop growing seasons pivotal to reproduce global fluctuations in maize and wheat yields. Sci. Adv. 4. https://doi.org/10.1126/sciadv.aat4517

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C.H., Rosenzweig, C., Wheeler, T.R., 2017. Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. Agric. Syst. 155, 269–288. https://doi.org/10.1016/j.agsy.2016.09.021

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. Eur. J. Agron. 18, 235–265. https://doi.org/10.1016/S1161-0301(02)00107-7

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron. 18, 267–288. https://doi.org/10.1016/S1161-0301(02)00108-9

Kimball, B.A., Boote, K.J., Hatfield, J.L., Ahuja, L.R., Stockle, C., Archontoulis, S., Baron, C., Basso, B., Bertuzzi, P., Constantin, J., Deryng, D., Dumont, B., Durand, J.L., Ewert, F., Gaiser, T., Gayler, S., Hoffmann, M.P., Jiang, Q., Kim, S.H., Lizaso, J., Moulin, S., Nendel, C., Parker, P., Palosuo, T., Priesack, E., Qi, Z., Srivastava, A., Stella, T., Tao, F., Thorp, K.R., Timlin, D., Twine, T.E., Webber, H., Willaume, M., Williams, K., 2019. Simulation of maize evapotranspiration: An inter-comparison among 29 maize models. Agric. For. Meteorol. 271, 264–284. https://doi.org/10.1016/j.agrformet.2019.02.037

Lange, S., 2019. EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI) [WWW Document]. GFZ Data Serv. https://doi.org/10.5880/pik.2016.004

Li, Y., Guan, K., Schnitkey, G.D., DeLucia, E., Peng, B., 2019. Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. Glob. Chang. Biol. 25, gcb.14628. https://doi.org/10.1111/gcb.14628

Lindeskog, M., Arneth, A., Bondeau, A., Waha, K., Seaquist, J., Olin, S., Smith, B., 2013. Implications of accounting for land use in simulations of ecosystem carbon cycling in Africa. Earth Syst. Dyn. 4, 385–407. https://doi.org/10.5194/esd-4-385-2013

Liu, J., Williams, J.R., Zehnder, A.J.B., Yang, H., 2007. GEPIC - modelling wheat yield and crop water productivity with high resolution on a global scale. Agric. Syst. 94, 478–493. https://doi.org/10.1016/j.agsy.2006.11.019

Liu, W., Yang, H., Folberth, C., Wang, X., Luo, Q., Schulin, R., 2016a. Global investigation of impacts of PET methods on simulating crop-water relations for maize. Agric. For. Meteorol. 221, 164–175. https://doi.org/10.1016/j.agrformet.2016.02.017

Liu, W., Yang, H., Liu, J., Azevedo, L.B., Wang, X., Xu, Z., Abbaspour, K.C., Schulin, R., 2016b. Global assessment of nitrogen losses and trade-offs with yields from major crop cultivations. Sci. Total Environ. 572, 526–537. https://doi.org/10.1016/j.scitotenv.2016.08.093

Minoli, S., Müller, C., Elliott, J., Ruane, A.C., Jägermeyr, J., Zabel, F., Dury, M., Folberth, C., François, L., Hank, T., Jacquemin, I., Liu, W., Olin, S., Pugh, T.A.M., 2019. Global Response Patterns of Major Rainfed Crops to Adaptation by Maintaining Current Growing Periods and Irrigation. Earth's Futur. 7, 1464–1480. https://doi.org/10.1029/2018EF001130

Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Folberth, C., Glotter, M., Hoek, S., Iizumi, T., Izaurralde, R.C., Jones, C., Khabarov, N., Lawrence, P., Liu, W., Olin, S., Pugh, T.A.M., Ray, D.K., Reddy, A., Rosenzweig, C., Ruane, A.C., Sakurai, G., Schmid, E., Skalsky, R., Song, C.X., Wang, X., De Wit, A., Yang, H., 2017. Global gridded crop model evaluation: Benchmarking, skills,

deficiencies and implications. Geosci. Model Dev. 10, 1403–1422. https://doi.org/10.5194/gmd-10-1403-2017
Müller, C., Elliott, J., Kelly, D., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Folberth, C., Hoek, S., Izaurralde, R.C., Jones, C.D., Khabarov, N., Lawrence, P., Liu, W., Olin, S., Pugh, T.A.M., Reddy, A., Rosenzweig, C., Ruane, A.C., Sakurai, G., Schmid, E., Skalsky, R., Wang, X., de Wit, A., Yang, H., 2019. The Global Gridded Crop Model Intercomparison phase 1 simulation dataset. Sci. Data 6, 50. https://doi.org/10.1038/s41597-019-0023-8
Müller, C., Robertson, R.D., 2014. Projecting future crop productivity for global economic modeling. Agric. Econ. 45, 37–50. https://doi.org/10.1111/agec.12088
Okada, M., Iizumi, T., Sakurai, G., Hanasaki, N., Sakai, T., Okamoto, K., Yokozawa, M., 2015. Modeling irrigation-based climate change adaptation in agriculture: Model development and evaluation in Northeast China. J. Adv. Model. Earth Syst. 7, 1409–1424. https://doi.org/10.1002/2014MS000402
Olin, S., Schurgers, G., Lindeskog, M., Wårlind, D., Smith, B., Bodin, P., Holmér, J., Arneth, A., 2015. Modelling the response of yields and tissue C : N to changes in atmospheric CO 2 and N management in the main wheat regions of western Europe. Biogeosciences 12, 2489–2515. https://doi.org/10.5194/bg-12-2489-2015
Porwollik, V., Müller, C., Elliott, J., Chryssanthacopoulos, J., Iizumi, T., Ray, D.K., Ruane, A.C., Arneth, A., Balkovič, J., Ciais, P., Deryng, D., Folberth, C., Izaurralde, R.C., Jones, C.D., Khabarov, N., Lawrence, P.J., Liu, W., Pugh, T.A.M., Reddy, A., Sakurai, G., Schmid, E., Wang, X., de Wit, A., Wu, X., 2017. Spatial and temporal uncertainty of crop yield aggregations. Eur. J. Agron. 88, 10–21. https://doi.org/10.1016/j.eja.2016.08.006
Ray, D.K., Gerber, J.S., Macdonald, G.K., West, P.C., 2015. Climate variation explains a third of global crop yield variability. Nat. Commun. 6, 1–9. https://doi.org/10.1038/ncomms6989
Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A.C.A.C., Müller, C., Arneth, A., Boote, K.J.K.J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T.A.M.T.A.M., Schmid, E., Stehfest, E., Yang, H., Jones, J.W.J.W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proc. Natl. Acad. Sci. U. S. A. 111, 3268–3273. https://doi.org/10.1073/pnas.1222463110
Ruane, A.C., Phillips, M.M., Rosenzweig, C., 2018. Climate shifts within major agricultural seasons for +1.5 and +2.0 °C worlds: HAPPI projections and AgMIP modeling scenarios. Agric. For. Meteorol. 259, 329–344. https://doi.org/10.1016/j.agrformet.2018.05.013
Sakurai, G., Iizumi, T., Nishimori, M., Yokozawa, M., 2014. How much has the increase in atmospheric CO2 directly affected past soybean production? Sci. Rep. 4, 1–5. https://doi.org/10.1038/srep04978
Schaphoff, S., Forkel, M., Müller, C., Knauer, J., von Bloh, W., Gerten, D., Jägermeyr, J., Lucht, W., Rammig, A., Thonicke, K., Waha, K., 2018. LPJmL4 – a dynamic global vegetation model with managed land – Part 2: Model evaluation. Geosci. Model Dev. 11, 1377–1403. https://doi.org/10.5194/gmd-11-1377-2018
Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. https://doi.org/10.1175/BAMS-D-11-00094.1
van Diepen, C.A., Wolf, J., van Keulen, H., Rappoldt, C., 1989. WOFOST: a simulation model of crop production. Soil Use Manag. 5, 16–24. https://doi.org/10.1111/j.1475-2743.1989.tb00755.x
van Ittersum, M.K., Leffelaar, P.A., van Keulen, H., Kropff, M.J., Bastiaans, L., Goudriaan, J., 2003. On approaches and applications of the Wageningen crop models. Eur. J. Agron. 18, 201–234. https://doi.org/10.1016/S1161-0301(02)00106-5
Wang, E., Martre, P., Zhao, Z., Ewert, F., Maiorano, A., Rötter, R.P., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Aggarwal, P.K., Anothai, J., Basso, B., Biernath, C., Cammarano, D., Challinor, A.J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A.-K., Liu, L., Müller, C., Naresh Kumar, S., Nendel, C., O'Leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ripoche, D., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P., Waha, K., Wallach, D., Wang, Z., Wolf, J., Zhu, Y., Asseng, S., Asseng, S., 2017. The uncertainty of crop yield projections is reduced by improved temperature response functions. Nat. Plants 3, 17102. https://doi.org/10.1038/nplants.2017.102
Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., Schewe, J., 2014. The inter-sectoral impact model intercomparison project (ISI-MIP): Project framework. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.1312330110
Williams, J.R., 1995. The EPIC model. Comput. Model. watershed Hydrol. 909–1000.
Williams, J.R., Jones, C.A., Dyke, P.T., 1984. A Modeling Approach to Determining the Relationship Between Erosion and Soil Productivity. Trans. ASAE 27, 0129–0144. https://doi.org/10.13031/2013.32748

Wu, X., Vuichard, N., Ciais, P., Viovy, N., de Noblet-Ducoudré, N., Wang, X., Magliulo, V., Wattenbach, M., Vitale, L., Di Tommasi, P., Moors, E.J., Jans, W., Elbers, J., Ceschia, E., Tallec, T., Bernhofer, C., Grünwald, T., Moureaux, C., Manise, T., Ligne, A., Cellier, P., Loubet, B., Larmanou, E., Ripoche, D., 2016. ORCHIDEE-CROP (v0), a new process-based agro-land surface model: model description and evaluation over Europe. Geosci. Model Dev. 9, 857–873. https://doi.org/10.5194/gmd-9-857-2016

You, L., Wood-Sichra, U., Fritz, S., Guo, Z., See, L., Koo, J., 2014. Spatial Production Allocation Model (SPAM) 2005 v2.0. MapSPAM [WWW Document]. URL http: //mapspam.info

Zhu, P., Zhuang, Q., Archontoulis, S. V., Bernacchi, C., Müller, C., 2019. Dissecting the nonlinear response of maize yield to high temperature stress with model-data integration. Glob. Chang. Biol. 25, 2470–2484. https://doi.org/10.1111/gcb.14632