









## The optimization of model ensemble composition and size can enhance the robustness of crop yield projections

Linchao Li <sup>1,2,3</sup>, Bin Wang <sup>3✉</sup>, Puyu Feng <sup>4</sup>, Jonas Jägermeyr<sup>5,6,7</sup>, Senthold Asseng <sup>8</sup>, Christoph Müller <sup>7</sup>, Ian Macadam<sup>9</sup>, De Li Liu <sup>3,10✉</sup>, Cathy Waters<sup>11</sup>, Yajie Zhang<sup>2</sup>, Qinsi He<sup>12</sup>, Yu Shi <sup>1,2</sup>, Shang Chen<sup>13</sup>, Xiaowei Guo<sup>14</sup>, Yi Li<sup>15</sup>, Jianqiang He<sup>15</sup>, Hao Feng<sup>1,15</sup>, Guijun Yang<sup>16</sup>, Hanqin Tian<sup>17</sup> & Qiang Yu <sup>1,2✉</sup>

Linked climate and crop simulation models are widely used to assess the impact of climate change on agriculture. However, it is unclear how ensemble configurations (model composition and size) influence crop yield projections and uncertainty. Here, we investigate the influences of ensemble configurations on crop yield projections and modeling uncertainty from Global Gridded Crop Models and Global Climate Models under future climate change. We performed a cluster analysis to identify distinct groups of ensemble members based on their projected outcomes, revealing unique patterns in crop yield projections and corresponding uncertainty levels, particularly for wheat and soybean. Furthermore, our findings suggest that approximately six Global Gridded Crop Models and 10 Global Climate Models are sufficient to capture modeling uncertainty, while a cluster-based selection of 3-4 Global Gridded Crop Models effectively represents the full ensemble. The contribution of individual Global Gridded Crop Models to overall uncertainty varies depending on region and crop type, emphasizing the importance of considering the impact of specific models when selecting models for local-scale applications. Our results emphasize the importance of model composition and ensemble size in identifying the primary sources of uncertainty in crop yield projections, offering valuable guidance for optimizing ensemble configurations in climate-crop modeling studies tailored to specific applications.

<sup>1</sup> College of Soil and Water Conservation Science and Engineering, Northwest A&F University, 712100 Yangling, China. <sup>2</sup> State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Northwest A&F University, 712100 Yangling, China. <sup>3</sup> NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW 2650, Australia. <sup>4</sup> College of Land Science and Technology, China Agricultural University, 100193 Beijing, China. <sup>5</sup> NASA Goddard Institute for Space Studies, New York, NY 10025, USA. <sup>6</sup> Columbia University, Climate School, New York, NY 10025, USA. <sup>7</sup> Potsdam Institute for Climate Impact Research (PIK), Leibniz Association, 14412 Potsdam, Germany. <sup>8</sup> Department of Life Science Engineering, Technical University of Munich, 85354 Freising, Germany. <sup>9</sup> CSIRO Climate Innovation, Hub, Australia. <sup>10</sup> Climate Change Research Centre, University of New South Wales, Sydney, NSW 2052, Australia. <sup>11</sup> NSW Department of Primary Industries, Dubbo, NSW 2830, Australia. <sup>12</sup> School of Life Sciences, Faculty of Science, University of Technology Sydney, PO Box 123, Broadway, NSW 2007, Australia. <sup>13</sup> Key Laboratory of Meteorological Disaster, Ministry of Education and Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science and Technology, Jiangsu Province, 210044 Nanjing, China. <sup>14</sup> Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, China. <sup>15</sup> College of Water Resources and Architectural Engineering, Northwest A&F University, 712100 Yangling, China. <sup>16</sup> Key Laboratory of Quantitative Remote Sensing in Agriculture of Ministry of Agriculture and Rural Affairs, Information Technology Research Center, Beijing Academy of Agriculture and Forestry Sciences, 100097 Beijing, China. <sup>17</sup> Schiller Institute for Integrated Science and Society, Department of Earth and Environmental Sciences, Boston College, Chestnut Hill, MA 02467, USA. ✉email: [bin.a.wang@dpi.nsw.gov.au](mailto:bin.a.wang@dpi.nsw.gov.au); [de.li.liu@dpi.nsw.gov.au](mailto:de.li.liu@dpi.nsw.gov.au); [yuq@nwfau.edu.cn](mailto:yuq@nwfau.edu.cn)

Climate change associated with increasing atmospheric greenhouse gas concentrations affects many agricultural regions and threatens global crop production and food security<sup>1–4</sup>. In recent years, climate change has increased the risks of simultaneous failure of major crops for breadbaskets across the globe<sup>5</sup> and projected climate change may push one-third of food production areas outside a “safe” climatic space<sup>6</sup>. Coincident with these factors, increasing population and food demand increase the difficulty of achieving the UN’s Sustainable Development Goal target of hunger eradication by 2030<sup>7,8</sup> and other health and environmental targets<sup>9</sup>.

Process-based biophysical crop models driven by climate data downscaled from multiple global climate models (GCMs) are often used to project crop yield changes and aid strategic decision-making for agriculture<sup>10–12</sup>. The Agricultural Model Intercomparison and Improvement Project (AgMIP)’s Global Gridded Crop Model Intercomparison (GGCMI) provides the essential data products needed to analyze global crop yield in response to climate change at global scales<sup>13–15</sup>. Projections based on the latest generation of GCMs, those contributing to the Coupled Model Intercomparison Project phase 6 (CMIP6)<sup>16</sup>, show climate change having a largely negative impact on crop yield in many breadbasket regions but also indicate some positive changes for individual crops, especially across higher latitudes<sup>3</sup>. However, there are large uncertainties in GCM–GGCM yield projections, both for the breadbasket regions<sup>2,17</sup> and at the global gridded scale<sup>18,19</sup>. These include modeling uncertainty related to the structure and parameters of the GCMs and crop models<sup>20–22</sup>. Recent analyses of uncertainty due to GCM and GGCM at a global scale for different crops have been conducted<sup>3,23</sup>. These studies reported that GGCMs are often the dominant source of uncertainty in quantifying relative changes in crop yield, even though the GGCM-induced variance decreased in the latest CMIP6-based projections<sup>3</sup> compared with an earlier GGCM ensemble-driven with CMIP5 climate projections<sup>18</sup>. Note that Müller et al.<sup>23</sup> also found that the overall variance of a crop-model emulator ensemble increases when driven by CMIP6 in comparison with being driven by the previous CMIP5 generation of GCM climate projections, but the relative shares of GGCM-induced variance were similar between the CMIP5 and CMIP6 ensembles. Wang et al.<sup>22</sup> found that the dominant source of uncertainty in wheat (*Triticum aestivum* L.) yield change was site-specific. A more comprehensive analysis of the sources of uncertainty at a gridded scale for four crops with a very large ensemble also showed that the exclusion of specific model combinations could lead to substantial differences in the representation of uncertainty<sup>23</sup>, raising the question of how much of the actual uncertainty can be represented by large ensembles.

Optimizing the sample size for GCM–GGCM projections depends on the level of uncertainty and the specific research question. However, the computational requirements of process-based crop models constrain the number of GCMs that can be used in harmonized crop model protocols, thereby influencing the uncertainty of yield projections<sup>24</sup>. The latest CMIP6-based GGCM Intercomparison Project (GGCMI) phase 3 projections provide the largest archive of process-based future crop yield projections as of today, with 60 combinations of 12 GGCMs and five GCMs<sup>3</sup>. The selection of five GCMs was a tradeoff between computational requirements and the representation of climate model uncertainty. Running GGCMs for additional GCMs would allow for more robust uncertainty attribution. However, the wide variation among the pool of models can obscure important information when relying solely on the ensemble mean. The effective use of large ensembles has received limited attention.

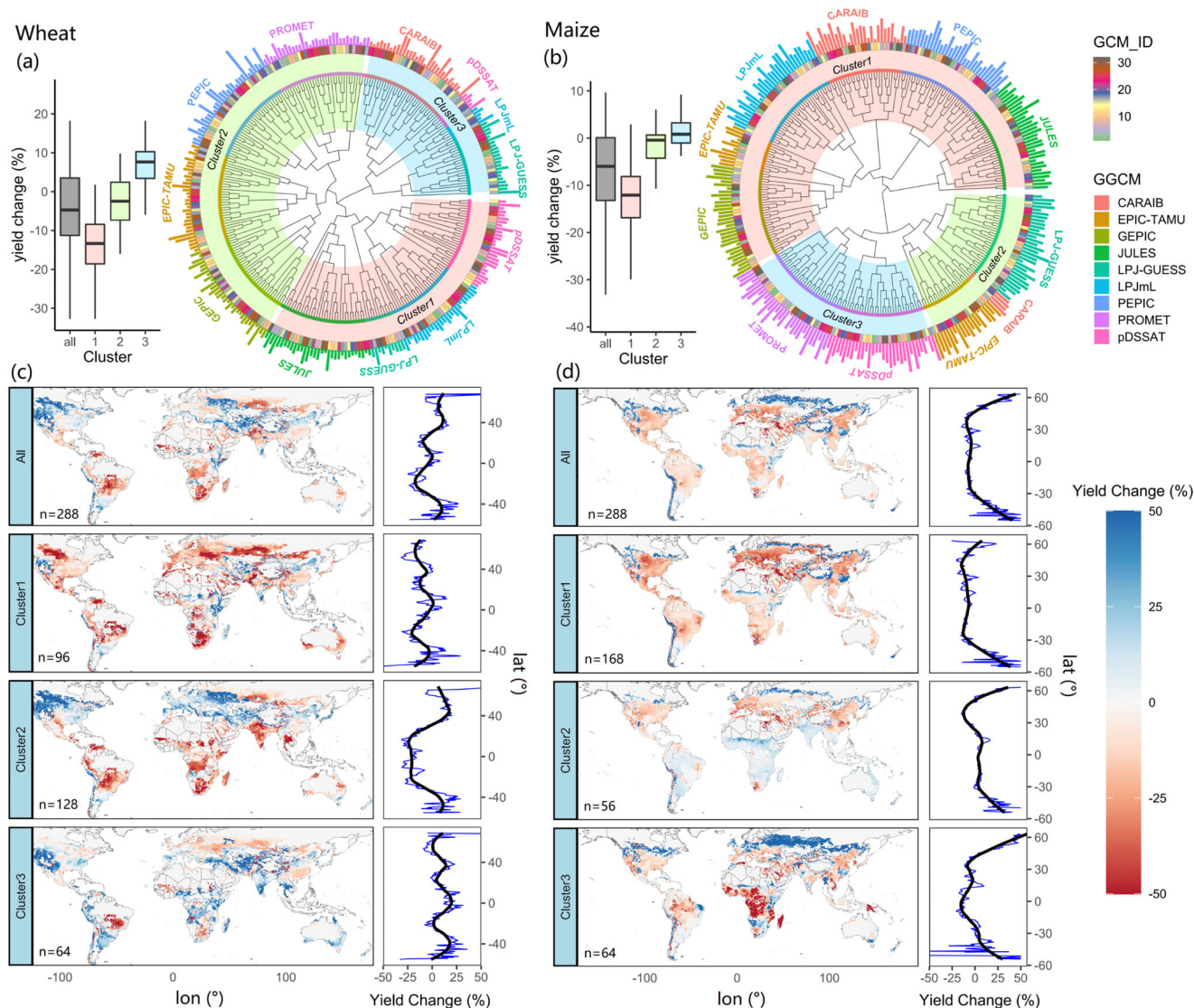
Several studies have established the minimum necessary ensemble size to effectively capture the uncertainty of models in

different fields, with the aim of reducing redundancies and offering guidance for model selection<sup>21,24–26</sup>. For instance, Falconnier, et al.<sup>27</sup> found that crop model ensemble skill increased with more maize (*Zea mays* L.) models considered in an ensemble in Sub-Saharan Africa. They reported that at least eight randomly selected models can ensure satisfactory accuracy in reproducing historical crop yield. However, this does not mean that this number of models is adequate for sampling uncertainty of future yield changes. Crop models that produce similar results for a historical time period can produce divergent results under future climate due to differences in model processes (e.g., temperature response, CO<sub>2</sub> fertilization effects)<sup>22</sup>. McSweeney and Jones<sup>28</sup> found that in order to cover the full spectrum of climate change projections for all regions, at least 13 GCMs needed to be considered. Ruane and McDermid<sup>29</sup> found that using five GCMs with different basic classes (e.g., relatively cool/wet, cool/dry, middle, hot/wet, and hot/dry) could capture the uncertainty across the full ensemble of simulated future climate change at a study site. However, the guidance for creating representative subsets to accurately reflect the uncertainty of global scale crop yield projections is still not well defined. The uncertainty of these projections is highly sensitive to the selection of crop and climate models<sup>23</sup>, and there has not yet been a comprehensive study to quantify the impact of ensemble size (GGCM and GCM) selected through random or science-based methods on uncertainty in the agricultural sector.

Wheat, maize, rice (*Oryza sativa* L.), and soybean (*Glycine max* (L.) Merr.) provide two-thirds of the caloric intake of the human population<sup>30</sup>. In this study, we present a new framework for better designing ensembles and extracting more information from a large pool of models. The framework utilizes nine GGCM emulators driven by changes in growing-season climate from 32 CMIP6 GCMs<sup>31</sup> and 12 GGCM phase 3 simulations driven by five GCMs<sup>3</sup>. Because our focus was on modeling uncertainties in the response of yields to climate change, we generated data for a single high greenhouse gas emissions scenario (SSP585) for the end of the 21st century. The study investigates the influences of ensemble configurations on crop yield projections and modeling uncertainty from GGCMs and GCMs under future climate change. We analyze the importance of model composition and ensemble size in determining the primary sources of uncertainty in crop yield projections, with the aim of offering valuable guidance for optimizing ensemble configurations in climate-crop modeling studies. We expect our analysis to enhance the accuracy of crop yield projections from a large ensemble, bridging the gap of notable ensemble uses and contributing to global food security.

## Results

**Yield projections by different ensembles.** We explored changes in simulated crop yield across GCMs for wheat, maize, rice, and soybean between 1980–2010 and 2069–2099 for each GGCM emulator and the multi-model ensemble mean (Fig. S1). Results were generated by considering the CO<sub>2</sub> fertilization effect. Without the CO<sub>2</sub> fertilization effect, most GCM–GGCM combinations showed a decline in simulated yield, although one or two GGCMs simulated increases in maize, rice, and soybean yields. As for the CO<sub>2</sub> fertilization effect, the wheat, rice, and soybean yield changes showed a stronger decrease signal compared with maize. In addition, crop yield changes with the CO<sub>2</sub> fertilization effect had a larger difference among the models. Note that the uncertainty in the CO<sub>2</sub> fertilization effect as implemented in the GGCMs is one of the major sources of the variance within the GGCM ensemble<sup>23</sup>. For instance, a large response to CO<sub>2</sub> fertilization was observed in the emulator of the JULES model,



**Fig. 1** Hierarchical clustering of wheat and maize yield changes (between 2069–2099 and 1980–2010) under SSP585. The left panels of **a** and **b** are boxplots showing the change in global yield for each cluster and all models. Box boundaries indicate the 25th and 75th percentiles of crop yield change. The black line within each box indicates the median value. The right panels of **a** and **b** show (from outside to inside) the GGCM names, the simulated yields of each GGCM (bar plots), the different GCMs, the three clusters, and the cluster tree diagrams, respectively. The three clusters are highlighted with light red, green, and blue. The left panels of **c** and **d** are the average wheat and maize yield changes for each cluster. The numbers ( $n=$ ) at the bottom left of each panel represent the number of ensemble members (GGCM  $\times$  GCM) used for each cluster. The right panels of **c** and **d** are the yield changes averaged by latitude (blue line) and the smoothed relationship between latitude and yield change (black line). Note that the wheat yield is the combined spring wheat and winter wheat yield based on harvested area data. GGCM is global gridded crop model; GCM is global climate model.

while the emulator of EPIC-TAMU showed relatively weaker CO<sub>2</sub> fertilization effects, especially for wheat and rice.

We used agglomerative-hierarchical clustering analysis (see the “Materials and methods” section) to divide yield projections (considering CO<sub>2</sub> fertilization) from GCM–GGCM combinations into three clusters (clusters 1–3) on the basis of spatial patterns and magnitudes of percentage yield changes and ranked yield changes of cluster 1 to cluster 3 from most negative to most positive for the four respective crops (Figs. 1a and b, and S2a, b, respectively). The clustering tended to group results by GGCM, indicating that, in general, differences between GGCMs were responsible for differences between clusters. For instance, cluster 2 in Fig. 1a illustrates that EPIC-TAMU, GEPIC, PEPIC, and PROMET (around 44%) are clustered together due to similar yield changes, yet they include the projections from all 32 GCMs. This result suggests that in this cluster, the crop models are the

primary factors influencing the projections of crop yield changes. However, it is worth noting that despite this crop model dominance in distinguishing ensemble members, different GCMs also contributed to the diversity of projections within each crop model group. This is evident as the associated GCMs in clusters 1 and 3 of Fig. 1a do not consistently group by crop models. For example, both cluster 1 and cluster 3 contain projections from pDSSAT, LPJ-GUESS, and LPJmL, each driven by different GCMs. Since the spatial patterns of yield changes varied between clusters, the yield changes for cluster 1 (3) were not always ranked the most negative (positive) in a particular region of the world (Fig. 1c, d). For instance, we can see positive wheat yield responses in some regions (e.g., East Africa, south India) in cluster 1 and these regions had some negative yield changes in cluster 2 (Fig. 1c). In addition, some regions experienced yield losses in the future regardless of clusters, including clusters with a



tendency towards increases in simulated yield. For example, wheat yield decreased in southern Brazil and southern Africa across three clusters (Fig. 1c), indicating that climate change will have negative effects on those areas.

We also generated cluster tree diagrams for the GGCM phase 3 simulation, which was improved with better inputs and an updated model version (Figs. S3–S6). The spatial patterns of yield changes among the different clusters displayed large variations, similar to those in the GGCM emulator. However, the ensemble member distance in GGCM phase 3 was distinct from that of the GGCM emulator. For instance, in Fig. 1a, the ensemble members (GGCMs driven by GCMs) are mainly grouped by the GGCMs, with only pDSSAT, LPJ-GUESS, and LPJmL appearing concurrently in clusters 1 and 3. For GGCM phase 3, the wheat yield changes within cluster 1 were primarily grouped by two specific GCMs (IPSL-CM6A-LR and UKESM1-0-LL). This can be attributed to the fact that these two GCMs exhibit a higher equilibrium climate sensitivity (ECS)<sup>32,33</sup>, which leads to expected large yield losses by some climate-sensitive crop models. However, it is worth noting that some crop models (e.g., DNDC, PROMET, and LPJmL) still fall in the same branch among the five GCMs (Fig. S3), indicating that different crop models have different levels of sensitivity to climate change.

We propose a novel approach for analyzing ensemble data by defining a distance metric between different ensemble members and reducing the large ensembles to three key patterns. This approach offers a deeper understanding of design ensemble configuration than relying solely on multi-model ensembles that may ultimately obscure results and overlook important details. Furthermore, our cluster scheme can potentially improve the harmonization of crop models and GCMs. For example, by identifying that sensitive crop models (e.g., those sensitive to warming) are driven by high ECS GCMs, we can anticipate that these models will experience a greater yield decline than other crop models (e.g., cluster 1 in Figs. S3 and S6).

**The uncertainty of different ensembles.** Crop yield simulations are impacted by local environment and management practices (e.g., climate conditions and crop management) that further influence the source of uncertainty in yield change projections. Therefore, we used such information as inputs for the cluster analysis (another type of cluster analysis for classifying regions, see details in the “Materials and methods” section) to divide the global cropping region into twelve sub-regions for each crop (Fig. S7).

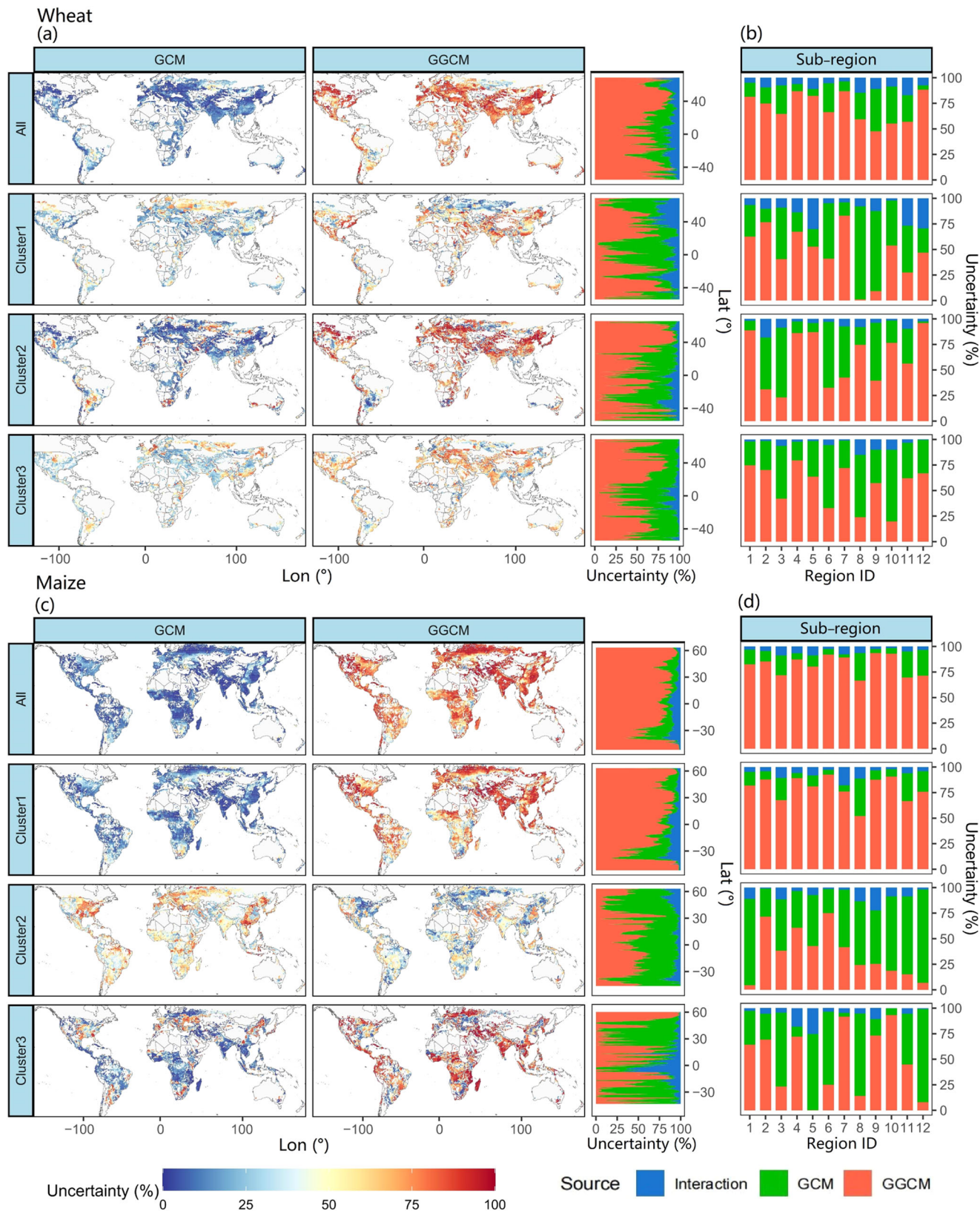
We explored the major source of modeling uncertainty for each subset (the combination of GGCM × GCM based on cluster analysis) for gridded, latitudinal, and regional scales (Fig. 2). For the entire ensemble of all GCMs and GGCMs, the crop model-induced variance share was larger than that of the GCMs (Fig. 2a, c, left panels). For different clusters, the dominant source of uncertainty for yield changes varied among different clusters, especially for maize, rice, and soybean (Figure S8). These results provide examples of how different model combinations can impact the sources of modeling uncertainty. For example, the ensemble size of cluster 2 ( $n = 56$  and  $n_{\text{GGCM}} = 3$ , Fig. 1d) and cluster 3 ( $n = 64$  and  $n_{\text{GGCM}} = 2$ , Fig. 1d) for maize were similar, but the dominant source of uncertainty exhibited large differences at different spatial scales (Fig. 2c, d). For instance, GCM was the dominant source of uncertainty in Northern Europe, Eastern Asia, and the USA for cluster 2 (Fig. 2c). However, in cluster 3, GGCM was the dominant source of uncertainty in these regions. In addition, the ensemble size of cluster 2 ( $n = 72$  and  $n_{\text{GGCM}} = 4$ , Fig. S2) and cluster 3 ( $n = 85$  and  $n_{\text{GGCM}} = 3$ , Fig. S2) for rice had a similar number of GGCMs, but they show differences in the

sources of uncertainties across sub-regions (Fig. S8). This is because, compared with cluster 3, the three GGCMs of cluster 2 had similar physiological processes (note that both EPIC-TAMU and GEPIC models were developed from EPIC). Several studies have investigated the source of uncertainty based on different sets of crop models<sup>19,22,34,35</sup>. However, the impact of model compositions, which can significantly influence the uncertainty of crop yield projections, has often been overlooked. For example, some studies that focused on the same region and used a similar number of crop models have shown varied results regarding the source of uncertainty<sup>35,36</sup>. Thus, we highlight the need to carefully select ensemble members in agricultural impact studies, given that different model compositions can lead to varying sources of uncertainty. In our study, although each cluster consisting of varying numbers of models, represents a unique potential scenario of yield changes under climate change. Such approaches can offer a more comprehensive understanding of the complexity and uncertainties involved in projecting crop yields under climate change.

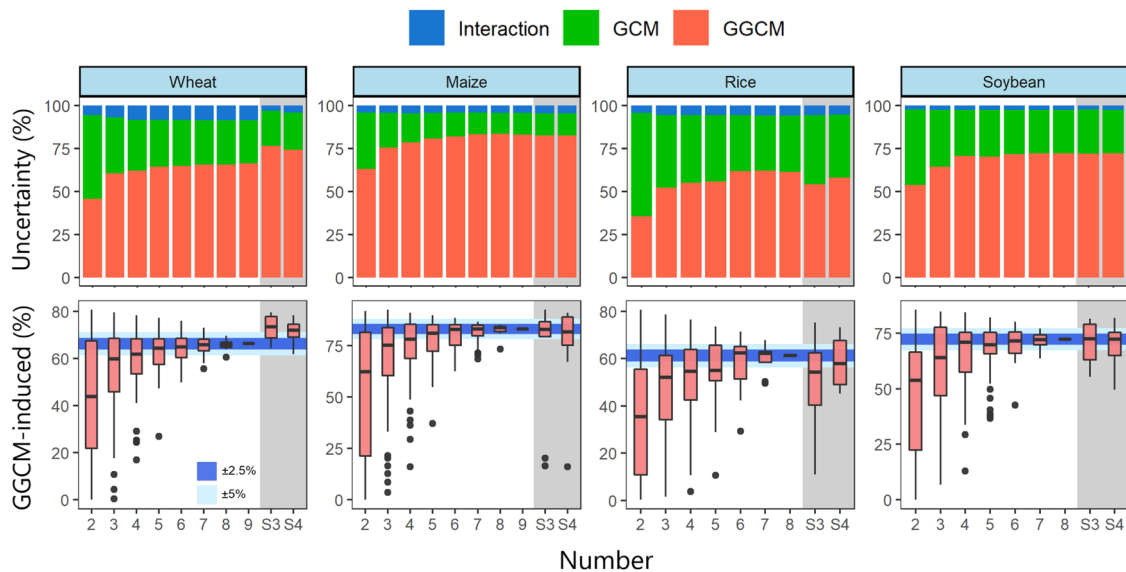
**Impacts of model quantity on uncertainty.** In order to reduce the quantity of simulations, some studies have ranked model performance and selected representative models for further analysis<sup>37–39</sup>, even though the selected models may not reflect the overall characteristics of the full ensemble. Here we quantify the uncertainty associated with the number of models, including nine GGCM emulators and 32 GCMs (Figs. 3 and S9). The results indicated that the GGCM- or GCM-induced uncertainty proportion first increased with the number of models (GGCM or GCM) used, and then became stable when the number of models exceeded a certain value. We found that for GGCM-emulators, randomly choosing at least five to six crop models (six GGCMs for wheat, maize, and rice, five GGCMs for soybean) and 9–12 GCMs (nine GCMs for wheat, 10 GCMs for maize, 12 GCMs for rice and soybean) adequately captured the uncertainty of yield changes compared with the full model set. For GGCM phase 3 simulations, there were six to seven GGCMs (seven GGCMs for wheat and soybean, six GGCMs for maize and rice) that could reflect the overall uncertainty of yield change projections. The least ensemble size varied between regions (Figs. S10, S11 and Tables S3–S4) but was within the range of four to six GGCMs and 6–14 GCMs for most regions. Also, the spatial patterns of GGCM-induced uncertainty and associated standard deviations changed little when the number of models was greater than approximately five (Figs. S12, S13).

One question that arises is whether the minimum effective ensemble size can be further reduced through the use of specific ensemble strategies. To address this, we used the family tree of models to select a subset of GGCMs, ensuring that at least one model belonged to each cluster. Our results indicated that for both GGCM-emulators and Phase 3 simulations, using 3–4 GGCMs was sufficient to capture the overall variance induced by the GGCMs (Figs. 3 and S14). However, this may not be the case for wheat yield projections in Phase 3 simulations (Fig. S14). The cluster tree for wheat yield (GGCM phase 3) did not always center around the GGCM, as demonstrated by cluster 1 of Fig. S3 which was mainly grouped by IPSL-CM6A-LR and UKESM1-0-LL. Generally, although our results indicated that a subset of GGCMs was sufficient to capture the overall variance, caution should still be exercised when using multi-model ensemble projections directly.

**Proportion of uncertainty attributed to individual GGCMs.** GGCMs were the dominant source of uncertainty in most of our study areas. Therefore, we here mainly focus on the uncertainty



**Fig. 2** The proportion of uncertainty from GCM and GGCM in simulated wheat and maize yield changes in 2069–2099 under SSP585 compared with 1980–2010. The left panels of **a** and **c** are the spatial patterns of uncertainty proportions from GCMs and GGCMs at the gridded scale across the globe for different clusters (see Fig. 1) and all ensemble members for wheat and maize. The right panels of **a** and **c** show the proportion of uncertainty at different latitudes for wheat and maize. To obtain the uncertainty at different latitudes, we first calculated the yield change averaged by latitude for each ensemble member and then quantified the proportion of uncertainty across different latitudes using ANOVA. **b** and **d** show the sources of uncertainty at different sub-regions (see Fig. S7). GGCM is the global gridded crop model; GCM is the global climate model.



**Fig. 3** The proportion of uncertainty of yield changes (2069–2099 compared with 1980–2010, under SSP585) for different numbers of GGCMs used for wheat, maize, rice, and soybean. The numbers 2–9 (and 2–8) represent different numbers of GGCMs used, including all of the possible combinations for wheat and maize (rice and soybean). The boxplot shows the distribution of uncertainty due to GGCM selection under different numbers of GGCMs used. The dark and light blue shaded areas respectively represent the thresholds of  $\pm 2.5\%$  and  $\pm 5\%$  when using the contribution of uncertainty estimated from all GGCMs as the benchmark. Box boundaries indicate the 25th and 75th percentiles. The black line within each box indicates the median value. GGCM is a global gridded crop model; GCM is a global climate model. S3 is the scenario when one model was randomly selected from each cluster (total 3 crop models). S4 is the scenario when four models were selected ensuring that at least one model belonged to each cluster.

attributed to individual GGCMs. We tested the proportion of uncertainty attributed to individual models by consecutively removing each crop model used, and then comparing the difference of GGCM-induced uncertainty between the full model set and without the particular model (Figs. 4 and S15–S17). For wheat, the CARAIB, PEPIC, and pDSSAT models added uncertainty in most regions, while the contributions of LPJ-GUESS and JULES varied among different regions. For example, the LPJ-GUESS model increased the GGCM-induced uncertainty share for wheat yield projections in Australia and decreased the GGCM-induced uncertainty share in southeast Asia. This variable result could be because LPJ-GUESS often is the outlier in strong  $\text{CO}_2$ -fertilization responses<sup>31</sup>. PEPIC and, especially, pDSSAT increased the GGCM-induced uncertainty share for maize yield projections. This could be due to pDSSAT having a strong temperature response for maize<sup>31</sup>. JULES increased the GGCM-induced uncertainty share in the Southern Hemisphere and decreased the GGCM-induced uncertainty share mainly in the Northern Hemisphere (Fig. S15). For rice, PEPIC mainly contributed to GGCM-induced uncertainty shares in most regions, especially in Africa; JULES increased the GGCM-induced uncertainty share in Asia; PROMET and pDSSAT had large influences on GGCM-induced uncertainty shares, but the direction varied among different regions (Fig. S16). For soybean, JULES increased the GGCM-induced uncertainty share in almost all regions; CARAIB, PROMET, and pDSSAT also greatly influenced the GGCM-induced uncertainty share, but this varied between different regions (Fig. S17). These results demonstrated that each GGCM affected the uncertainty of model predictions across the globe to some degree, highlighting the importance of model selection at the local scale.

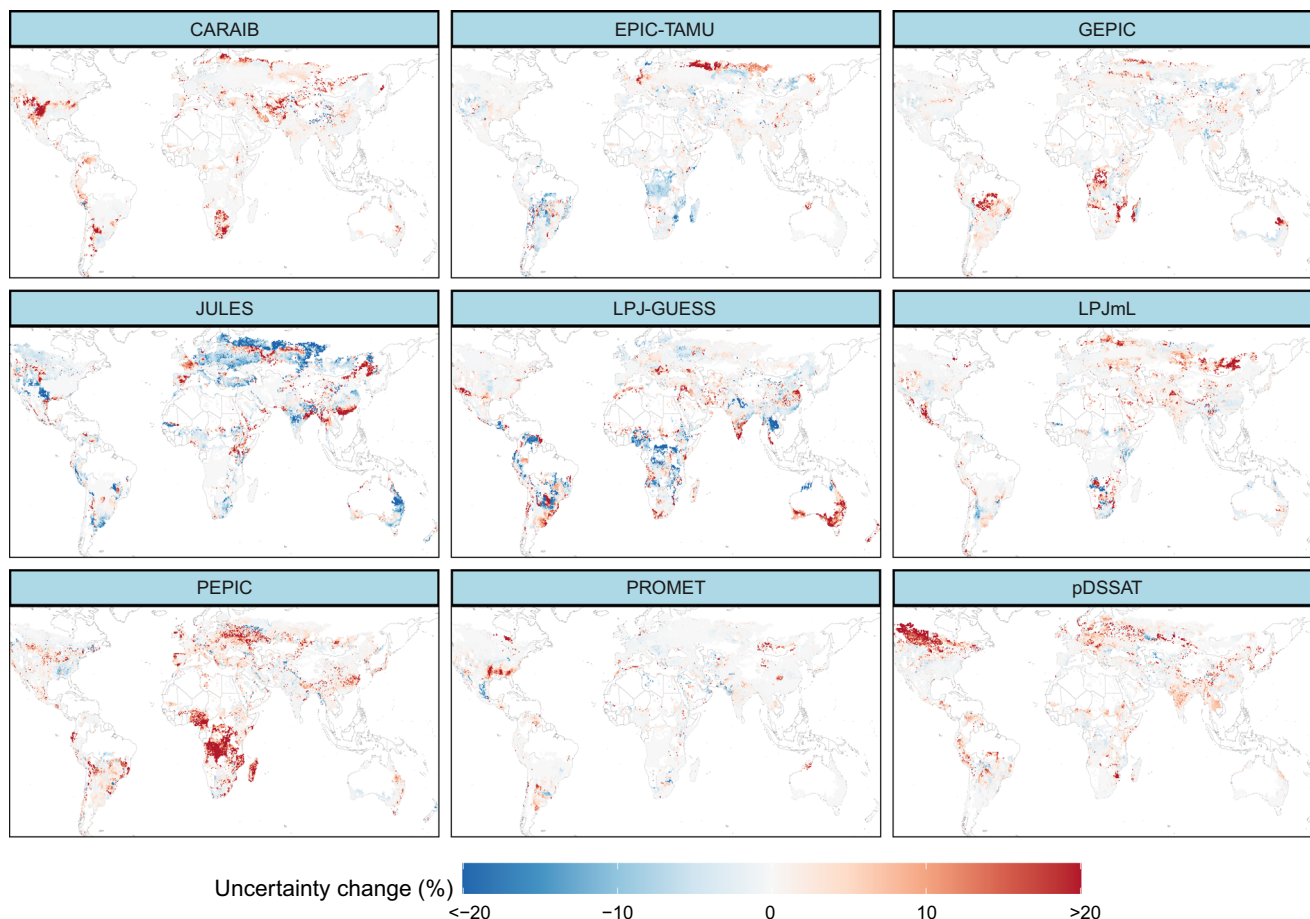
**Implications and limitations.** Although the multi-model ensemble mean is a common summary statistic for yield projections, it has limited utility for determining adaptation strategies that are robust to a range of plausible future yield trajectories, and

it may give a misleading result of little change where individual model combinations have changes of different signs. Thus, it is necessary to consider results from diverse collections of models. Our results can provide information for carefully designing GCM–GGCM ensembles that maximize the sampling of uncertainty. We focused on the SSP585 emissions trajectory to explore yield changes under a large climate change signal. Müller et al.<sup>23</sup> found that the relative importance of different sources of uncertainty in yield projections was comparable between SSP126 and SSP585 for these GGCM emulators, while overall cross-ensemble variance was substantially larger under SSP585. Therefore, the results provided here are likely representative of other levels of climate change as well.

Understanding the inter-model distances and relationships using hierarchical clustering and family trees offers valuable insights into the similarities and differences among ensemble members, supporting the improvement and diversification of crop models and their responses to climate change<sup>40,41</sup>. For instance, by identifying and addressing any dependence among models, the overall uncertainty in the ensemble can be more robustly estimated, leading to more reliable yield projections. In addition, our study investigated the influence of ensemble size on uncertainties. We found that the number of models had a large influence on the represented uncertainty in different regions across the globe. For example, using 10 GCMs captured the GCM-induced uncertainty for maize yield change at the global scale (Fig. S9), but more GCMs were needed for the USA, Central Asia, and the southern part of South America (sub-regions 1, 3, 8, and 11) (Fig. S11 and Table S4).

The different yield changes for different ensemble members have important implications for the design of ensemble configurations (Figs. 1 and S2–S6). For example, the JULES simulations of soybean yields were distinct from results obtained with other models and made a large contribution to overall uncertainty, especially in South Asia and Central Europe. How such extreme behavior is dealt with may have a large effect on the results from the multi-model ensemble and corresponding





**Fig. 4** Changes in the GCM-induced uncertainty share contributed by individual GCMs at the grid level for wheat in 2069–2099 under SSP585. The contribution was determined by sequentially removing one of the nine GCMs from the ensemble of GCMs at each grid point, and then comparing the GCM-induced uncertainty share of projected yield changes between the full GCM model set and the set without the particular model. A positive change indicates that including a particular model increased the GCM-induced uncertainty share. A negative contribution change indicates that using this model decreased the GCM-induced uncertainty share. GCM is global gridded crop model.

uncertainty<sup>42</sup>. In these kinds of situations, some previous studies have excluded or down-weighted a given model from the ensemble in order to reduce the overall uncertainty<sup>22,43</sup>. Therefore, it is very important to carefully screen and evaluate the performance of the individual models for their robustness to simulate climate change impacts on crop yield before they are selected for use in a multi-model ensemble for the purpose of reducing uncertainty in climate change impact assessment<sup>44</sup>. Additionally, sensitive crop models driven by high ECS GCMs may also lead to extreme behavior (Figs. S3–S6), while some GCMs may not have significant influence from such high ECS GCMs. These results highlight the ongoing challenge of designing an optimal model ensemble. Our study proposes dividing ensemble members (crop models driven by GCMs) into groups based on their characteristics in order to provide a more effective way to select the most suitable ensemble members for a specific application. This strategy can efficiently represent the full ensemble and reduce redundancy.

Different methods have been developed to improve crop model performance and reduce overall modeling uncertainty, such as improving model structure<sup>45,46</sup> or parameter optimization<sup>19,47</sup>. However, conducting this work at the global scale is challenging due to high computation costs and missing calibration target data. Recently, some studies have constrained the uncertainty of GCMs by using the results of field experiments and statistical models<sup>48,49</sup>. These studies have suggested that using more

detailed information from field experiments can improve model responses to climate warming and further reduce the overall uncertainty. Other factors such as CO<sub>2</sub> fertilization are also likely to impact crop growth and bring uncertainty to yield projections. Several studies found that the CO<sub>2</sub> fertilization effect was constrained by frequent extreme weather events<sup>50,51</sup>. However, such impacts have not been fully considered in current simulations, and consequently, these simulations have probably underestimated the yield loss and increased the overall uncertainty<sup>48</sup>. Therefore, to better understand the interacting impacts of climate change (e.g., heatwave, extreme precipitation, drought, and CO<sub>2</sub> fertilization) on crop yields and thereby reduce uncertainty in climate change impact assessment, more controlled environment experiments with long time periods and large ranges of crop genotypes in a wider range of mega-environments are needed<sup>52,53</sup>.

We recognize several limitations in our study. Some limitations are related to the GCMs. First, sources of uncertainty such as soil data and management options (e.g., fertilization rate)<sup>2,54–57</sup> and the effects of pests and diseases were not considered by the GCM simulations<sup>18</sup>. Second, model parameterization also influences the projections and affects uncertainty. For example, Xiong et al.<sup>19</sup> analyzed the uncertainty resulting from parameterization using four parameterization strategies. While structural differences between GCMs were accounted for in our study, uncertainty in parameter settings for individual GCMs was not

considered. Third, although some GGCMs can capture yield losses that occur under extreme temperature events (including heat and frost) and drought<sup>58</sup>, they underestimate the magnitude of yield losses due to extremely wet conditions<sup>17,59</sup>. This is compounded by a fourth limitation related to our use of the change factor method to generate future climate. This method does not fully sample the uncertainty due to changes in extreme climate events, or event changes related to within-season variability (emulators only driven by climate change during the crop growing season), as simulated by the GCMs. Furthermore, the emulator may not perfectly reproduce the raw GGCM simulation. Although we also used the GGCM phase 3 dataset that used bias-adjusted daily GCM data to compare dominant sources of uncertainty with GGCM emulators, only five GCMs were currently available. Greater efforts are needed to better harmonize the processes of climate and crop models as we look towards closing in on a “certain” projection from a large ensemble.

## Materials and methods

**GGCM emulators (GGCMI phase 2).** Franke et al.<sup>31</sup> developed emulators for many of the GGCMI phase 2 models to facilitate the use of numerous GCM–GGCM combinations to analyze climate change impacts on crop yields<sup>23</sup>. Crop model emulators have allowed larger sets of GCMs, for example, all CMIP6 GCMs, to be considered in assessments of uncertainties in yield predictions<sup>23</sup>. The emulators combined the advantages of both process-based and statistical crop models, and have already been widely used in previous studies<sup>60,61</sup>. The GGCM emulators can successfully capture crop yield responses to temperature, precipitation, CO<sub>2</sub> concentration, and nitrogen change. In addition, GGCM emulators have the potential to project crop yields under future climate change<sup>31</sup>. The change factor method was used to generate future climate data for GGCM emulators. This approach uses future monthly rainfall and temperature changes from the GCMs to perturb historical daily data without changing the distribution of rainfall and temperature. The GGCM Phase 2 models used in our study are not formally calibrated<sup>14,31</sup>. It is conceivable that a lack of calibration could influence the relative changes that are the basis of our study, but note that calibration can sometimes have negative impacts on model skill<sup>15</sup>.

We used the emulators to simulate crop yields at the grid scale for wheat, maize, rice, and soybean. Emulators for nine GGCMs (CARAIB, EPIC-TAMU, JULES, GEPIC, LPJ-GUESS, LPJmL, pDSSAT, PEPIC, and PROMET) for wheat and maize, and eight GGCMs (CARAIB, EPIC-TAMU, JULES, GEPIC, LPJmL, pDSSAT, PEPIC, and PROMET) for rice and soybean were available. The crop yield changes were simulated under both rainfed and irrigated conditions, and we aggregated the rainfed and irrigated yield based on the crop area data of Monfreda et al.<sup>62</sup> for maize, rice, and soybean. Winter wheat and spring wheat were simulated individually; however, the data of Monfreda et al.<sup>62</sup> did not distinguish the harvest areas of winter wheat and spring wheat. In this study, the harvest areas of winter wheat and spring wheat were distinguished by average temperatures and the length of the growing season according to Müller et al.<sup>23</sup>. Winter wheat yields and spring wheat yields were then combined.

The emulators used annual values of growing season mean temperature and total precipitation on a 0.5° × 0.5° global grid and atmospheric CO<sub>2</sub> concentrations as climate inputs. Because we did not consider growing season adaptation, only the A0 (no growing season adaptation) setting was used in this study. The nitrogen application rate for each grid was the same as used by Elliott et al.<sup>13</sup>. In this study, we mainly focused on the crop yield

change with CO<sub>2</sub> fertilization. More detailed information can be found in Müller et al.<sup>23</sup>. Nitrogen application rate and the areas of rainfed and irrigated crops were assumed to be the same for our baseline and future time periods (1980–2010 and 2069–2099). We focused on the relative yield change (%) for each grid under climate change for model comparisons<sup>18</sup>.

**GGCMI phase 3 crop modeling.** The GGCMI phase 3 project<sup>3</sup> involved five CMIP6 GCMs (GFDL-ESM4, IPSL-CM6A-LR, MPI-ESM1-2-HR, MRI-ESM2-0, and UKESM10-LL) and 12 GGCMs (ACEA, CROVER, CYGMA1p74, DSSAT-Pythia, EPIC-IIASA, ISAM, LandscapeDNDC, LPJmL, pDSSAT, PEPIC, PROMET, and SIMPLACE-LINTUL5). Daily GCM data were bias-adjusted and downscaled to a 0.5° × 0.5° global grid by the ISIMIP ([www.isimip.org](http://www.isimip.org)) framework<sup>63–65</sup> prior to input to the GGCMs. These GCMs were selected to best sample the full CMIP6 ensemble in terms of the sensitivity of global mean temperature to increasing atmospheric carbon dioxide concentrations (i.e., equilibrium climate sensitivity) and other factors.

In this study, each model within the ensemble has been subjected to a benchmark performance evaluation for the historical period, ensuring that the models meet a certain level of quality<sup>3,14,31</sup>. We use the relative change approach, which focuses on the differences between baseline and future projections, accounting for proportional changes instead of absolute values. This method allows for a more comprehensive evaluation of the influence of ensemble configurations, such as model composition and size, on crop yield projections and uncertainties under future climate change scenarios.

**Climate data.** We downloaded monthly temperature and precipitation data from simulations of 32 GCMs under SSP585 from the CMIP6 archives (Table S1, <https://esgf-node.llnl.gov/search/cmip6/>). We resampled this data to the 0.5° grid using a bilinear interpolation method. We used the change factor method to generate data for GGCM emulators. Since our study only focuses on the changes in future multi-year average temperature ( $\Delta T$ ) and precipitation ( $\Delta P$ ) in relation to the baseline of the same GCM, no bias correction is necessary for the computation of  $\Delta T$  and  $\Delta P$ . For each GCM, simulated differences in growing season mean temperature and growing season total precipitation between 1980–2010 and 2069–2099 were applied to the baseline data from the AgMIP Modern-Era Retrospective Analysis for Research and Applications (AgMERRA). The global average atmospheric CO<sub>2</sub> content was obtained from the International Institute for Applied Systems Analysis (<https://iiasa.ac.at/>). A CO<sub>2</sub> concentration of 810 ppm was used for SSP585 in 2069–2099, and the baseline concentration was 360 ppm. The baseline yield for each GCM was the same for a given GGCM because the change factor method used AgMERRA data for the baseline period.

## Two types of cluster analysis

**Cluster analysis of crop-growing regions.** We divided the global cropping regions into different sub-regions based on the characteristics of local crop productivity, climate conditions, and management<sup>10</sup>. Four kinds of environmental variables with the  $k$ -mean cluster analysis method were used to determine different sub-regions for wheat, maize, rice, and soybean (Table S2). To reflect the crop yield during the baseline period (1980–2010), we used ensembles of either nine or eight GGCMs to determine the average crop yield during the historical period. We selected mean temperature, solar radiation, precipitation, air relative humidity, and wind speed as the main climate factors for the  $k$ -mean analysis (Table S2). Temperature, water, and light are the main factors influencing the speed and rate of germination, emergence,



and growth of crops<sup>66</sup>. The climate data we used were from AgMERRA, and are the same as the emulators' inputs<sup>31</sup>. We also used latitude and longitude to represent the location information. Nitrogen application rate was also considered as a management option based on Müller et al.<sup>23</sup>. It is worth noting that the GGCM-emulators take into account management practices including water supply and nitrogen inputs. Given that precipitation is separately accounted for in the cluster analysis, within the framework of the GGCM-emulator projections, only nitrogen fertilizer input was represented as a distinct management variable. Finally, 12 sub-regions were determined for each crop.

Currently, there are several regionalizing plans available, including the IPCC's climate zones<sup>67</sup>, global agroecological zones (AEZs), and global food production units (FPUs). Despite their existence, these plans may not be able to effectively capture the regional diversity of sources of uncertainty in crop yields. This is because of the differences in harvest area distributions and factors that influence crop yield.

**Cluster analysis of future yield change.** We used agglomerative-hierarchical clustering analysis<sup>41</sup> to reveal the different ensemble configurations of the relative yield changes when CO<sub>2</sub> fertilization was considered. The Canberra distance (*D*) was used to compute the initial cluster distance among ensemble members<sup>68,69</sup>. This approach was performed downstream (without any preselection), using the results of crop yield changes at every grid location. This allowed us to cluster ensemble members with high similarity in terms of spatial pattern and magnitude<sup>41</sup>. There were 288 members (9 GGCMs × 32 GCMs) for wheat and maize, and 256 members (8 GGCMs × 32 GCMs) for rice and soybean. We classified future yield change into three clusters to show three considerably dissimilar yield change predictions. Cluster 1 represented yield decreases, cluster 2 represented small changes in yield, and cluster 3 represented yield increases. Since crop yield is a common and key variable in crop modeling, the investigation of different variables might result in only minor variations in the outcomes.

### Uncertainty analysis

**ANOVA methodology.** We used the analysis of variance (ANOVA) technique to quantify the uncertainty in crop yield changes for wheat, maize, rice, and soybean due to GGCMs and GCMs. The equations for the total sum of squares (SST) for a two-way ANOVA are:

$$SST = SS_{GGCM} + SS_{GCM} + SS_{GGCM * GCM} \quad (1)$$

$$P_{GGCM}(\%) = SS_{GGCM}/SST \times 100 \quad (2)$$

$$P_{GCM}(\%) = SS_{GCM}/SST \times 100 \quad (3)$$

$$P_{GGCM * GCM}(\%) = SS_{GGCM * GCM}/SST \times 100 \quad (4)$$

where  $P_{GGCM}$  is the uncertainty contribution due to GGCM,  $P_{GCM}$  is the uncertainty contribution due to GCM, and  $P_{GGCM * GCM}$  is the uncertainty contribution arising from interactive effects between GCMs and GGCMs. In this study, we use the mean proportion of uncertainty as it offers a straightforward and readily interpretable measure of central tendency.

**Subsampling scheme.** We used ANOVA to estimate how different sample sizes of models would affect the total uncertainty in yield change. For GGCMs, different combinations were selected as  $C_i^9$  for wheat and maize,  $C_i^8$  for rice and soybean,  $i$  varies from 2, 3,

...,9 (or 8).

$$\overline{SST}_{(i)} = \frac{1}{n} \sum_{j=1}^n (SS_{GGCM(i)} + SS_{GCM} + SS_{GGCM(i) \times GCM}) \quad (5)$$

$$P_{GGCM(i)}(\%) = \frac{\frac{1}{n} \sum_{j=1}^n (SS_{GGCM(i)})}{\overline{SST}_{(i)}} \times 100 \quad (6)$$

where  $i$  is the number of the GGCM used (2–9 for wheat and maize, 2–8 for rice and soybean),  $n$  is the total number of  $C_i^9$  for wheat and maize,  $C_i^8$  for rice and soybean,  $j$  is 1, 2, ...,  $n$ .

The ANOVA analysis for the number of GCMs was similar to the analysis for the number of GGCMs. However, because the number of GCMs was 32, the combinations of  $C_i^{32}$  would be more than thousands of millions when  $i$  was between 6 and 26. It was therefore not practical to use all combinations when subsampling GCMs. Consequently, we tested the uncertainty change under different sizes of random subsets and found that the uncertainty change tended to be stable when the size was around 100,000 ( $i$  ranging from six to 27) for each crop (Fig. S18). Therefore, we set the maximum size of subsampling GCM to be 100,000.

$$\overline{SST}_{(i)} = \begin{cases} \frac{1}{n} \sum_{j=1}^n (SS_{GGCM} + SS_{GCM(i)} + SS_{GGCM \times GCM(i)}) & 2 \leq i \leq 5 \text{ or } i \geq 28 \\ \frac{1}{100,000} \sum_{j=1}^{100,000} (SS_{GGCM} + SS_{GCM(i)} + SS_{GGCM \times GCM(i)}) & 6 \leq i \leq 27 \end{cases} \quad (7)$$

$$P_{GCM(i)}(\%) = \begin{cases} \frac{\frac{1}{n} \sum_{j=1}^n (SS_{GCM(i)})}{\overline{SST}_{(i)}} \times 100 & 2 \leq i \leq 5 \text{ or } i \geq 28 \\ \frac{\frac{1}{100,000} \sum_{j=1}^{100,000} (SS_{GCM(i)})}{\overline{SST}_{(i)}} \times 100 & 6 \leq i \leq 27 \end{cases} \quad (8)$$

where  $i$  is the number of GCMs used,  $n$  is the total number of  $C_i^{32}$ ,  $j$  ranges from 1 to  $n$

In this study, we defined the least number of GGCMs necessary to sample uncertainty to be when the GGCM-induced uncertainty was within  $\pm 2.5\%$  of the uncertainty for the full ensemble of GGCMs<sup>24</sup>. However, because the proportion of GCM-induced uncertainty was lower than that for GGCM, we used a threshold of  $\pm 1.5\%$  to select the minimum ensemble size for GCMs.

### Data availability

All data required to evaluate the conclusions in this paper can be found in the paper itself or the Supplementary Materials. The CMIP6 data can be accessed at <https://esgf-node.llnl.gov/projects/cmip6/>. Publicly available model inputs are obtainable from the ISMIP website (<https://www.isimip.org/>) or by contacting the corresponding author. The GGCM crop calendar can be accessed at <https://doi.org/10.5281/zenodo.5062513>, and fertilizer input data is available at <https://doi.org/10.5281/zenodo.4954582>. Crop model emulators can be found at <https://doi.org/10.5281/zenodo.3592453> and <https://doi.org/10.5281/zenodo.4321276>.

### Code availability

The code of the cluster analysis can be found in the "cluster" v2.1.4R package that is available at <https://www.rdocumentation.org/packages/cluster>. The cluster tree diagrams are made by the "ggtree" v1.4.11R package that is available at <https://rdocumentation.org/packages/ggtree>. The code for crop model emulators can be found at Müller et al. (2021) (<https://zenodo.org/record/4321276>). The detailed code for data processing and illustration is available from the corresponding author upon reasonable request.

Received: 3 May 2023; Accepted: 15 September 2023;

Published online: 09 October 2023

### References

1. Mbwo, C. et al. *Climate Change and Land: an IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management*,

- Food Security and Greenhouse Gas Fluxes in Terrestrial Ecosystems* (IPCC, 2019).
2. Hasegawa, T. et al. Extreme climate events increase risk of global food insecurity and adaptation needs. *Nat. Food* **2**, 587–595 (2021).
  3. Jägermeyr, J. et al. Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nat. Food* <https://doi.org/10.1038/s43016-021-00400-y> (2021).
  4. Lesk, C. et al. Stronger temperature–moisture couplings exacerbate the impact of climate warming on global crop yields. *Nat. Food* **2**, 683–691, (2021).
  5. Gaupp, F., Hall, J., Hochrainer-Stigler, S. & Dadson, S. Changing risks of simultaneous global breadbasket failure. *Nat. Clim. Change* **10**, 54–57 (2019).
  6. Kummu, M., Heino, M., Taka, M., Varis, O. & Viviroli, D. Climate change risks pushing one-third of global food production outside the safe climatic space. *One Earth* **4**, 720–729 (2021).
  7. FAO. *World Food and Agriculture—Statistical Pocketbook* (FAO, 2018).
  8. Molotoks, A., Smith, P. & Dawson, T. P. Impacts of land use, population, and climate change on global food security. *Food Energy Secur.* **10**, <https://doi.org/10.1002/fes3.261> (2020).
  9. Bodirsky, B. L. et al. The ongoing nutrition transition thwarts long-term targets for food security, public health and environmental protection. *Sci. Rep.* **10**, 19778 (2020).
  10. Chapagain, R. et al. Decomposing crop model uncertainty: a systematic review. *Field Crops Res.* **279**, <https://doi.org/10.1016/j.fcr.2022.108448> (2022).
  11. Wang, B. et al. Australian wheat production expected to decrease by the late 21st century. *Global Change Biol.* **24**, 2403–2415 (2018).
  12. Asseng, S. et al. Rising temperatures reduce global wheat production. *Nat. Clim. Change* **5**, 143–147 (2014).
  13. Elliott, J. et al. The Global Gridded Crop Model Intercomparison: data and modeling protocols for Phase 1 (v1.0). *Geosci. Model Dev.* **8**, 261–277 (2015).
  14. Franke, J. A. et al. The GGCM Phase 2 experiment: global gridded crop model simulations under uniform changes in CO<sub>2</sub>, temperature, water, and nitrogen levels (protocol version 1.0). *Geosci. Model Dev.* **13**, 2315–2336 (2020).
  15. Müller, C. et al. Global gridded crop model evaluation: benchmarking, skills, deficiencies and implications. *Geosci. Model Dev.* **10**, 1403–1422 (2017).
  16. Eyring, V. et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016).
  17. Li, Y., Guan, K., Schnitkey, G. D., DeLucia, E. & Peng, B. Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. *Glob. Chang. Biol.* **25**, 2325–2337 (2019).
  18. Rosenzweig, C. et al. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proc. Natl Acad. Sci. USA* **111**, 3268–3273 (2014).
  19. Xiong, W. et al. Different uncertainty distribution between high and low latitudes in modelling warming impacts on wheat. *Nat. Food* **1**, 63–69 (2019).
  20. Tao, F. et al. Contribution of crop model structure, parameters and climate projections to uncertainty in climate change impact assessments. *Glob. Chang. Biol.* **24**, 1291–1307 (2018).
  21. Asseng, S. et al. Uncertainty in simulating wheat yields under climate change. *Nat. Clim. Change* **3**, 827–832 (2013).
  22. Wang, B. et al. Sources of uncertainty for wheat yield projections under future climate are site-specific. *Nat. Food* **1**, 720–728 (2020).
  23. Müller, C. et al. Exploring uncertainties in global crop yield projections in a large ensemble of crop models and CMIP5 and CMIP6 climate scenarios. *Environ. Res. Lett.* **16**, <https://doi.org/10.1088/1748-9326/abd8fc> (2021).
  24. Wang, H. M., Chen, J., Xu, C. Y., Zhang, J. & Chen, H. A framework to quantify the uncertainty contribution of GCMs over multiple sources in hydrological impacts of climate change. *Earth's Future* **8**, <https://doi.org/10.1029/2020ef001602> (2020).
  25. Solazzo, E. & Galmarini, S. A science-based use of ensembles of opportunities for assessment and scenario studies. *Atmos. Chem. Phys.* **15**, 2535–2544 (2015).
  26. Lee, J. et al. Robust evaluation of ENSO in climate models: how many ensemble members are needed? *Geophys. Res. Lett.* **48**, <https://doi.org/10.1029/2021gl095041> (2021).
  27. Falconnier, G. N. et al. Modelling climate change impacts on maize yields under low nitrogen input conditions in sub-Saharan Africa. *Glob. Chang. Biol.* **26**, 5942–5964 (2020).
  28. McSweeney, C. F. & Jones, R. G. How representative is the spread of climate projections from the 5 CMIP5 GCMs used in ISI-MIP? *Clim. Serv.* **1**, 24–29 (2016).
  29. Ruane, A. C. & McDermid, S. P. Selection of a representative subset of global climate models that captures the profile of regional changes for integrated climate impacts assessment. *Earth Perspect.* **4**, <https://doi.org/10.1186/s40322-017-0036-4> (2017).
  30. FAO. United Nation's Food and Agricultural Organization. *Food Security Indicators*. <http://www.fao.org/faostat/> (2019).
  31. Franke, J. A. et al. The GGCM Phase-2 emulators: global gridded crop model responses to changes in CO<sub>2</sub>, temperature, water, and nitrogen (version 1.0). *Geosci. Model Dev.* **13**, 3995–4018 (2020).
  32. Bonnet, R. et al. Increased risk of near term global warming due to a recent AMOC weakening. *Nat. Commun.* **12**, 6108 (2021).
  33. Galik, C. S. A continuing need to revisit BECCS and its potential. *Nat. Clim. Change* **10**, 2–3 (2019).
  34. Jiang, T. et al. Identifying sources of uncertainty in wheat production projections with consideration of crop climatic suitability under future climate. *Agric. For. Meteorol.* **319**, <https://doi.org/10.1016/j.agrformet.2022.108933> (2022).
  35. Huang, M. et al. Dominant sources of uncertainty in simulating maize adaptation under future climate scenarios in China. *Agric. Syst.* **199**, <https://doi.org/10.1016/j.agry.2022.103411> (2022).
  36. Zhang, Y., Zhao, Y. & Feng, L. Higher contributions of uncertainty from global climate models than crop models in maize-yield simulations under climate change. *Meteorol. Appl.* **26**, 74–82 (2019).
  37. Li, L. et al. Future projections of extreme temperature events in different sub-regions of China. *Atmos. Res.* **217**, 150–164 (2019).
  38. Rattalino Edreira, J. I. et al. Spatial frameworks for robust estimation of yield gaps. *Nat. Food* **2**, 773–779 (2021).
  39. Wang, B. et al. Multi-model ensemble projections of future extreme temperature change using a statistical downscaling method in south eastern Australia. *Clim. Change* **138**, 85–98 (2016).
  40. Abramowitz, G. et al. ESD Reviews: Model dependence in multi-model climate ensembles: weighting, sub-selection and out-of-sample testing. *Earth Syst. Dyn.* **10**, 91–105 (2019).
  41. Morim, J. et al. Robustness and uncertainties in global multivariate wind-wave climate projections. *Nat. Clim. Change* **9**, 711–718 (2019).
  42. Knutti, R. The end of model democracy? *Clim. Change* **102**, 395–404 (2010).
  43. Vetter, T. et al. Evaluation of sources of uncertainty in projected hydrological changes under climate change in 12 large-scale river basins. *Clim. Change* **141**, 419–433 (2017).
  44. Wang, E. et al. The uncertainty of crop yield projections is reduced by improved temperature response functions. *Nat. Plants* **3**, 17102 (2017).
  45. Maiorano, A. et al. Crop model improvement reduces the uncertainty of the response to temperature of multi-model ensembles. *Field Crops Res.* **202**, 5–20 (2017).
  46. Liu, K. et al. Silver lining to a climate crisis in multiple prospects for alleviating crop waterlogging under future climates. *Nat. Commun.* **14**, 765 (2023).
  47. Wallach, D. et al. How well do crop modeling groups predict wheat phenology, given calibration data from the target population? *Eur. J. Agron.* **124**, 126195 (2021).
  48. Wang, X. et al. Emergent constraint on crop yield response to warmer temperature from field experiments. *Nat. Sustain.* **3**, 908–916 (2020).
  49. Zhao, C. et al. Plausible rice yield losses under future climate warming. *Nat. Plants* **3**, 16202 (2016).
  50. Gray, S. B. et al. Intensifying drought eliminates the expected benefits of elevated carbon dioxide for soybean. *Nat. Plants* **2**, 16132 (2016).
  51. Obermeier, W. A. et al. Reduced CO<sub>2</sub> fertilization effect in temperate C3 grasslands under more extreme weather conditions. *Nat. Clim. Change* **7**, 137–141 (2016).
  52. Kimball, B. A. Crop responses to elevated CO<sub>2</sub> and interactions with H<sub>2</sub>O, N, and temperature. *Curr. Opin. Plant Biol.* **31**, 36–43 (2016).
  53. Ottman, M., Kimball, B., White, J. & Wall, G. Wheat growth response to increased temperature from varied planting dates and supplemental infrared heating. *Agron. J.* **104**, 7–16 (2012).
  54. Folberth, C. et al. Parameterization-induced uncertainties and impacts of crop management harmonization in a global gridded crop model ensemble. *PLoS ONE* **14**, e0221862 (2019).
  55. McCullough, E. B., Quinn, J. D. & Simons, A. M. Profitability of climate-smart soil fertility investment varies widely across sub-Saharan Africa. *Nat. Food* **3**, 275–285 (2022).
  56. Folberth, C. et al. Uncertainty in soil data can outweigh climate impact signals in global crop yield simulations. *Nat. Commun.* **7**, 11872 (2016).
  57. Minoli, S., Jägermeyr, J., Asseng, S., Urfels, A. & Müller, C. Global crop yields can be lifted by timely adaptation of growing periods to climate change. *Nat. Commun.* **13**, 7079 (2022).
  58. Schauburger, B. et al. Consistent negative response of US crops to high temperatures in observations and crop models. *Nat. Commun.* **8**, 13931 (2017).
  59. Liu, W. et al. Heterogeneous impacts of excessive wetness on maize yields in China: Evidence from statistical yields and process-based crop models. *Agric. Forest Meteorol.* **327**, <https://doi.org/10.1016/j.agrformet.2022.109205> (2022).
  60. Liu, W. et al. Future climate change significantly alters interannual wheat yield variability over half of harvested areas. *Environ. Res. Lett.* **16**, <https://doi.org/10.1088/1748-9326/ac1fbb> (2021).

61. Zabel, F. et al. Large potential for crop production adaptation depends on available future varieties. *Global Change Biol.* <https://doi.org/10.1111/gcb.15649> (2021).
62. Monfreda, C., Ramankutty, N. & Foley, J. A. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochem. Cycles* **22**, n/a–n/a (2008).
63. Cucchi, M. et al. WFDE5: bias-adjusted ERA5 reanalysis data for impact studies. *Earth Syst. Sci. Data* **12**, 2097–2120 (2020).
64. Lange, S. Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geosci. Model Dev.* **12**, 3055–3070 (2019).
65. Lange, S. ISIMIP3BASD (2.5.0). *Zenodo*. <https://doi.org/10.5281/zenodo.4686991> (2021).
66. Monti, A. (ed.) *Switchgrass: A Valuable Biomass Crop For Energy* (Springer Science & Business Media, 2012).
67. Iturbide, M. et al. An update of IPCC climate reference regions for subcontinental analysis of climate model data: definition and aggregated datasets. *Earth Syst. Sci. Data* **12**, 2959–2970 (2020).
68. Becker, R., Chambers, J. & Wilks, A. *The New S Language. Computer Science Series* (Wadsworth & Brooks/Cole, Pacific Grove, CA, 1988).
69. Borg, I. & Groenen, P. J. *Modern Multidimensional Scaling: Theory and Applications* (Springer Science & Business Media, 2005).

## Acknowledgements

This study was supported by the collaboration project jointly founded by the NSFC (grant no. 41961124006) and the NSF (grant no. 1903722) for Innovations at the Nexus of Food, Energy, and Water Systems (INFEWS: US-China). We acknowledge the Agricultural Intercomparison and Improvement Project (AgMIP), the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP), and the contributing modelers for making crop model simulations available. We acknowledge the World Climate Research Program, which, through its Working Group on Coupled Modeling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies that support CMIP6 and ESGF. L.L. acknowledges a scholarship from the Chinese Scholarship Council and the NSW Department of Primary Industries provided office facilities for conducting this work.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s43247-023-01016-9>.

**Correspondence** and requests for materials should be addressed to Bin Wang, De Li Liu or Qiang Yu.

**Peer review information** *Communications Earth & Environment* thanks Fasil Rettie and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Heike Langenberg, Alienor Lavergne. A peer review file is available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023