Climate model uncertainty in impact assessments for agriculture: A multi-ensemble case study on maize in sub-Saharan Africa

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1002/2017EF000539">http://dx.doi.org/10.1002/2017EF000539</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Wiley Blackwell</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Fri Nov 23 00:36:40 EST 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/109497">http://hdl.handle.net/1721.1/109497</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution-NonCommercial-NoDerivs License</td>
</tr>
</tbody>
</table>
| Detailed Terms| http://creativecommons.org/licenses/by-nc-nd/4.0/
Climate model uncertainty in impact assessments for agriculture: A multi-ensemble case study on maize in sub-Saharan Africa

Amy Dale\(^1\), Charles Fant\(^1\), Kenneth Strzepek\(^1\), Megan Lickley\(^2\), and Susan Solomon\(^1\)\(^2\)

\(^1\)Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, \(^2\)Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Abstract We present maize production in sub-Saharan Africa as a case study in the exploration of how uncertainties in global climate change, as reflected in projections from a range of climate model ensembles, influence climate impact assessments for agriculture. The crop model AquaCrop-OS (Food and Agriculture Organization of the United Nations) was modified to run on a \(2^\circ \times 2^\circ\) grid and coupled to 122 climate model projections from multi-model ensembles for three emissions scenarios (Coupled Model Intercomparison Project Phase 3 [CMIP3] SRES A1B and CMIP5 Representative Concentration Pathway [RCP] scenarios 4.5 and 8.5) as well as two “within-model” ensembles (NCAR CCSM3 and ECHAM5/MPI-OM) designed to capture internal variability (i.e., uncertainty due to chaos in the climate system). In spite of high uncertainty, most notably in the high-producing semi-arid zones, we observed robust regional and sub-regional trends across all ensembles. In agreement with previous work, we project widespread yield losses in the Sahel region and Southern Africa, resilience in Central Africa, and sub-regional increases in East Africa and at the southern tip of the continent. Spatial patterns of yield losses corresponded with spatial patterns of aridity increases, which were explicitly evaluated. Internal variability was a major source of uncertainty in both within-model and between-model ensembles and explained the majority of the spatial distribution of uncertainty in yield projections. Projected climate change impacts on maize production in different regions and nations ranged from near-zero or positive (upper quartile estimates) to substantially negative (lower quartile estimates), highlighting a need for risk management strategies that are adaptive and robust to uncertainty.

1. Introduction

Uncertainty in projections of the socioeconomic, ecological, and health impacts of climate change has become a major topic of study in the last decade. Early impact studies relied on output from single general circulation model (GCM) runs; this is no longer an accepted method, and most authors now use multiple GCMs to generate a range of possible climate futures, thereby limiting the influence of errors in any one model [Tebaldi and Knutti, 2007; Hawkins and Sutton, 2011].

Multi-model GCM ensembles are necessary for the detailed exploration of uncertainty in climate model projections, but they are not necessarily sufficient. Often, each member of an ensemble is represented by a single model run, or one possible realization of a system whose behavior—due to the chaotic nature of the climate system—is highly sensitive to small perturbations of initial conditions. Thus, multi-model ensembles confound the large uncertainties that arise from sources of internal variability such as El Niño with structural uncertainties related to between-model differences in model structure, parameter selection, and boundary conditions (“model uncertainty”) [Tebaldi and Knutti, 2007]. The choice of greenhouse gas emission scenario also gives rise to “scenario uncertainty.”

The relative influence of each source of uncertainty varies with time; the influence of scenario uncertainty increases over time, whereas the influence of internal variability and model uncertainty decreases [Hawkins and Sutton, 2011]. Deser et al. [2012] attribute at least half of all uncertainty in the Coupled Model Intercomparison Project Phase 3 (CMIP3) multi-model ensemble from 2005 to 2060 to internal variability. At regional scales (~2,500 km), Hawkins and Sutton [2011] found that internal variability in projections of decadal mean
precipitation changes was comparable in magnitude to scenario uncertainty, even when the latter was at 2100 levels.

“Within-model” ensembles are comprised of multiple runs of a single GCM that differ only with respect to small variations in the initial conditions. Because within-model ensembles are subject only to internal variability, comparisons of “within-” and “between-model” ensembles shed important light on the magnitude of internal variability and its relative influence. Such comparisons have practical relevance in the development of appropriate adaptation and mitigation strategies: they separate the sources of uncertainty that are theoretically reducible (model and scenario uncertainty) from those that largely are not (internal variability) [Hawkins and Sutton, 2009].

Boehlert et al. [2015] used multiple within- and between-model GCM ensembles to assess the impact of climate model uncertainties on projected changes in runoff, crop water requirements, and adaptation costs out to 2090. Our work extends the multi-ensemble approach of Boehlert et al. [2015] to the problem of climate change impacts on maize production in sub-Saharan Africa (SSA).

Nations in SSA are expected to experience substantial future changes in agricultural production due to climate change [Niang et al., 2014] and have relatively limited capital and institutional capacity to address the potential threat [Challinor et al., 2007; Knox et al., 2012]. Almost a quarter of the world’s malnourished population lives in SSA [Lobell et al., 2008], yet, agriculture in this region comprises up to 50% of gross domestic product (GDP) in some countries [Schlenker and Lobell, 2010]. Maize is the most widely produced crop in SSA by harvested area [Food and Agriculture Organization of the United Nations, 2015] and is also the most calorically important [Schlenker and Lobell, 2010]. As a drought-sensitive species, it is also one of the crops that is most likely to be negatively impacted by climate change [Schlenker and Lobell, 2010; Rippke et al., 2016]. This is especially true in SSA, where the vast majority of cereal crop production relies on rainfall rather than irrigation [Challinor et al., 2007]. A systematic review of over 50 studies on climate change impacts on African crop yields from the field to regional scales revealed that the projected maize impacts of climate change display large uncertainties, with projected future changes ranging from nearly −50 to +100% [Knox et al., 2012].

Previous studies on maize production under climate change either globally or in SSA have relied on a single GCM or on limited multi-model ensembles [Lobell et al., 2008; Schlenker and Lobell, 2010; Thornton et al., 2011; Knox et al., 2012; Waha et al., 2013; Rosenzweig et al., 2014; Rippke et al., 2016]. A much richer range of within and between-model climate model ensembles has become available since these studies, and a deeper elucidation of uncertainty using this information is the primary focus and motivation for this paper.

We present results for a recently released open-source process-based crop model (FAO AquaCrop-OS v5.0a) coupled to projections from 122 GCM runs representing three between-model ensembles, two within-model ensembles, three emission scenarios, and both CMIP3 and CMIP5. This approach allowed us to identify regions in which the potential risks (and benefits) to the food supply due to climate change are robust across and within ensembles, as well as regions in which disagreement persists. It allowed us to evaluate the potential benefits of greenhouse gas mitigation, separate the influence of internal variability from that of model uncertainty and scenario uncertainty, and attribute regional shifts in crop production (and their associated uncertainties) to underlying shifts (and uncertainties) in temperature, precipitation, and aridity. We conclude with a comparison of our projections to projections from previous works and a discussion of their implications for the development of more robust impact assessments and management strategies.

2. Data and Methods

2.1. Climate Data and Processing
We used five ensembles from the World Climate Research Programme’s (WCRP’s) CMIP phase 3 (CMIP3) and phase 5 (CMIP5) multi-model dataset [Meehl et al., 2007; Taylor et al., 2012]. The three CMIP3 ensembles, which were run with the Special Report on Emissions Scenarios (SRES) A1B scenario, included a between-model ensemble of 22 projections from different GCMs and two within-model ensembles, 17 projections for the European Centre Hamburg Model Version 5/Max Planck Institute-Ocean Model (ECHAMS/MIP-OM) model [Sterl et al., 2008] and 40 projections for the National Center for Atmospheric Research Community Climate System Model Version 3 (NCAR CCSM3) [Deser et al., 2012]. The two CMIP5
ensembles included 23 projections for the Representative Concentration Pathway (RCP) 4.5 scenario (a mid-range emissions scenario) and 20 projections for RCP8.5 (high emissions scenario). The names of the specific CMIP3 and CMIP5 models used in this study are presented in Tables S1 and S2, Supporting Information [also presented in Boehlert et al., 2015]. GCM projections were regridded from their native resolution to the common 2° × 2° resolution as described by Boehlert [2015].

We consider four eras: “2010,” “2030,” “2050,” and “2090.” Each era represents an average over 20 years of projected changes in temperature and precipitation. For example, “2010” represents 2000–2019 and “2050” represents 2040–2059. Results are not available for the NCAR within-model ensemble beyond 2060.

In order to correct the biases in the GCMs, we applied the simple change factor method presented by Hawkins et al. [2013] (the “delta” method) to historical data from Princeton’s global meteorological forcing dataset [Sheffield et al., 2006]. This method is common in climate change impact studies [Arndt et al., 2011; Boehlert et al., 2015; Fant et al., 2016]. In this approach, we “nudge” historical daily climate data over the period 1961–1990 by changes in the mean temperature and precipitation projected for a given era by the GCMs. For example, for the 2050 era, the mean change in temperature (the “delta”) was taken to be the monthly mean temperature difference between GCM projections from 2040 to 2059 and the monthly mean from 1961 to 1990. This “delta” was then added to the daily historical climate data for the same baseline period, 1961–1990. In order to maintain the daily precipitation pattern, we applied the same method as for temperature but apportioned the monthly mean change in precipitation on a given day according to the fraction of the total monthly historical precipitation that occurred on that day. The crop yield reported below for each era is the average annual yield over a 30-year simulation (preceded by 5 years of model “spin up”) driven by the 1961–1990 climate nudged by computed changes in temperature and precipitation.

A possible shortcoming of the delta method is that it cannot address changes in climate variability over time. However, this assumption may be preferable in light of the ongoing uncertainties regarding the ability of GCMs to reliably project future changes in climate variability [Blöschl et al., 2007; Brown and Wilby, 2012].

2.2. Crop Model

AquaCrop (Food and Agriculture Organization of the United Nations) is a one-dimensional (“point”) model designed for use at the field scale. It performs well for maize under a wide range of environmental conditions and nutrient management scenarios [Heng et al., 2009; Hsiao et al., 2009; Steduto et al., 2011; Abedinpour et al., 2012; Kim and Kaluarachchi, 2015; Greaves and Wang, 2016], although its performance can suffer under severe water stress [Hsiao et al., 2009; Ahmadi et al., 2015; Greaves and Wang, 2016]. Despite its comparatively simple structure and limited input requirements, its performance was also comparable to, or superior to, that of other crop models in three cross-model comparisons for sunflower, barley, winter wheat, and maize [Todorovic et al., 2009; Eitzinger et al., 2013; Saab et al., 2015].

We refer readers to the AquaCrop manual for a detailed description of its formulation and data requirements [Raes et al., 2012]. Briefly, AquaCrop models the daily climate-driven soil water balance as well as crop growth, transpiration, and senescence in the presence or absence of management scenarios such as irrigation and fertilization. Five “water stress coefficients” are varied daily from 1 (no stress) to 0 (maximum stress) in order to modify crop-specific parameters that determine the rates of canopy expansion, transpiration, senescence, and root growth as well as the harvest index (HI, or the fraction of the total above-ground biomass that results in yield). Biomass and yield are calculated as follows:

\[ B = K_{sb} \cdot WP^* \sum_d \frac{Tr_d}{ET_{0d}} \]  

(1)

\[ Y = HI \cdot B \]  

(2)

where \( B \) is the total above-ground dry biomass over the entire growing season, \( K_{sb} \) is a temperature stress coefficient, \( WP^* \) is the crop water productivity (i.e., the crop’s ability to convert water consumed/transpired into biomass) normalized by the daily reference evapotranspiration and atmospheric \( CO_2 \), \( Tr_d \) is the crop transpiration on day \( d \); \( ET_{0d} \) is the reference evapotranspiration on day \( d \); and \( Y \) is the crop yield. Normalization of \( WP^* \) by \( ET_{0d} \) and \( CO_2 \) ensures that AquaCrop is applicable under a wide range of atmospheric conditions. As is suggested by the appearance of \( K_{sb} \) in the biomass formula, AquaCrop considers several
sources of stress in addition to water: heat stress, cold stress, soil infertility, and (not included in our work) soil salinity.

We used the Modified Hargreaves method adjusted for a daily timestep to estimate ET$_{0}$ [Farmer et al., 2011]. Droogers and Allen [2002] showed that Modified Hargreaves is likely to outperform the more standard and substantially more data-intensive Penman–Monteith method [Allen et al., 1998] when meteorological data are uncertain or error-prone, as is commonly the case for weather stations in SSA [Vrieling et al., 2013].

We modified FAO’s recently released open-source MATLAB version of AquaCrop (AquaCrop-OS v5.0a) to run for all grid cells in SSA at 2° × 2° resolution [Foster, 2016]. The conversion of site-scale crop models to gridded crop models in this manner is a common practice [Rosenzweig et al., 2014]. The supporting methods (including Figures S1–S6) provide details on model setup, calibration, and benchmarking procedures and results. Key model features are summarized here.

We modeled two soil layers with different hydraulic properties (topsoil and subsoil) using 18 soil compartments that expanded in thickness from 1 cm at the surface to 30 cm at depth. Pedotransfer functions [U.S. Department of the Interior, Bureau of Reclamation, 1993; Nemes et al., 2005; Mohamed and Ali, 2006] were used to estimate the hydraulic properties in each grid cell from the subsoil and topsoil composition reported in the Regridded Harmonized World Soil Database v1.2 [Wieder et al., 2014]. To ensure accurate initial conditions, we modeled the soil water balance for 5 years before collecting model results (“spin-up”). We did not model groundwater table effects on water logging and capillary rise [Fan et al., 2013].

In addition to climate-driven (rain-fed) crop growth, we accounted for the influence of fertilizer application [Stoorvogel et al., 1993; Sanchez, 2002] and irrigation [Frenken, 2005] on crop yields. However, fertilizer use, irrigation, and the maize harvested area were held constant over all eras in order to explore climate change impacts in the absence of land use change. Estimates of the harvested area in each grid cell and nitrogen, phosphorus, and potash fertilizer application rates for maize were downloaded from EarthStat (www .earthstat.org) [Monfreda et al., 2008; Mueller et al., 2012]. The ability of fertilizer to improve crop yields was captured as part of a multi-step calibration procedure based on AquaCrop’s soil fertility stress equations. The crop yields reported in the results are the weighted results for a “no irrigation” and “complete irrigation” scenario, where the weighting factor is the fraction of the harvested maize area in each grid cell that was equipped for irrigation [Frenken, 2005; Siebert et al., 2005; Siebert et al., 2013]. Raw fertilizer and irrigation data were upscaled from 0.05° × 0.05° to 2° × 2° as needed by taking the weighted average with respect to the fraction of the harvested maize area represented by each subcell.

We used a simplified version of the temperature and precipitation rules proposed by Waha et al. [2012] to optimize planting dates in each grid cell based on the climate for the current year. As a result, our projections may be more conservative than those of models that do not allow planting dates to adapt to climate change (e.g., four out of seven of the Agricultural Model Intercomparison and Improvement Project [AgMIP] models [Rosenzweig et al., 2014]).

We did not account for improved yields due to CO$_2$ enrichment. However, CO$_2$ effects are thought to be smaller for C$_4$ crops such as maize than they are for C$_3$ crops [Schlenker and Lobell, 2010]. Interestingly, two meta-analyses found no significant difference in projected crop yield responses to climate change between C$_3$ and C$_4$ crops, suggesting a potentially limited role for CO$_2$ effects in the face of other cross-study differences [Knox et al., 2012; Challinor et al., 2014].

We calibrated (“tuned”) the model by minimizing the root mean squared error between national yields reported by FAOSTAT for the year 2000 [Food and Agriculture Organization of the United Nations, 2015] and the 7-year average yields produced by the model when driven by historical Princeton climate data from 1997 to 2003. After calibration, we compared AquaCrop’s performance against that of three gridded crop models from the AgMIP Fast Track project (LPJmL, PEGASUS, and pDSSAT [Deryng et al., 2011; Rosenzweig et al., 2014]) with respect to both the FAOSTAT data and aggregated sub-national estimates from FAO [2016], Fischer et al. [2012], and Mueller et al. [2012] (Figures S5–S6). At 2° × 2° resolution, our model performed notably better than PEGASUS and pDSSAT and marginally better than LPJmL. This is to be expected, since our calibration procedure was designed to minimize error in projections for SSA rather than the entire globe.

We did not model any cell for which over 90% of the 0.05° × 0.05° subcells contained no harvest maize area [Monfreda et al., 2008]. Results are also not shown for any cell in which the simulation failed for over one
### Table 1. Color Rules for All Figures

<table>
<thead>
<tr>
<th>Category</th>
<th>Criteria</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 “No substantial change”</td>
<td>At least two thirds of the ensemble members resulted in a change between −5% and +5%</td>
<td>Gray</td>
</tr>
<tr>
<td>2 “Substantial but unknown”</td>
<td>Did not fall into Category 1 and less than two thirds of the ensemble members agreed on the direction of change</td>
<td>White</td>
</tr>
<tr>
<td>3 “Substantial and positive”</td>
<td>Did not fall into Category 1 or 2 and resulted in a positive change</td>
<td>Green</td>
</tr>
<tr>
<td>4 “Substantial and negative”</td>
<td>Did not fall into Category 1 or 2 and resulted in a negative change</td>
<td>Orange</td>
</tr>
</tbody>
</table>

third of the ensemble members due to inadequate conditions for crop growth. These cells are shown in dark gray in all figures.

### 3. Results

We used AquaCrop to estimate the annual maize yield and production (yield × harvested area) in every grid cell in Africa for each ensemble member, ensemble, and era. The figures that follow use a common coloring scheme to describe median projected percent changes in yields and underlying climate drivers (Table 1): grid cells for which at least two thirds of the ensemble members resulted in a change between −5% and +5% were colored light gray, indicating cells for which projections of small or no changes were robust across GCM ensemble members. Cells that did not fall into the previous category, and for which less than two thirds of the ensemble members agreed on the direction of change, were colored white; these are areas in which a substantial change could occur, but there is disagreement regarding the direction of change. Cells that did not meet the criteria for the first two categories are cells in which a robust change was observed across ensemble members (i.e., greater than two thirds of the ensemble members agree on the direction of change and less than two thirds suggest a change within 5%); these cells were colored green or orange depending on whether the direction of change was positive or negative.

We present medians rather than means because yield and production estimates cannot fall below zero (or −100% change); thus, the distribution of projections for any given ensemble tends to be skewed, especially in cells that experienced large negative impacts. Note that the ensemble median in all maps (e.g., Figure 1) was calculated cell-by-cell and thus represents a composite scenario generated from all ensemble members, rather than results for a single representative ensemble.

#### 3.1. Projected Climate Change

Figure 1 presents the median change in temperature, precipitation, and aridity index from 2010 to 2090 for the lowest and highest emission scenarios, CMIP5 RCP4.5 and RCP8.5. This figure lays the groundwork for a more detailed discussion of spatial variation in the projected maize yield response across all five ensembles (Section 3.2) and the relative influence of climate drivers on the yield response (Section 3.3).

The aridity index is a measure of the dryness of the terrestrial climate and was estimated as shown:

\[
AI = \frac{\sum P_d}{\sum ET_{od}}
\]  

where \(\sum P_d\) is the total annual precipitation (summed over all days \(d\)) and \(\sum ET_{od}\) is the total annual reference evapotranspiration [Middleton and Thomas, 1992; Fu and Feng, 2014]. Note that a positive percent change in \(AI\) implies decreasing aridity, or an increasingly wet climate.

As expected, both RCP4.5 and RCP8.5 project universally robust temperature increases, whereas within-ensemble disagreement regarding the direction of change in precipitation (white cells) is widespread, especially in RCP4.5. The RCP8.5 scenario projects robust increases in precipitation across much of the northern half of the continent and robust decreases south of the Democratic Republic of the Congo. Africa is projected to become more arid overall.

#### 3.2. Projected Yield Change

Figure 2 shows the median percent change in crop yields between 2010 and all other eras for each of the ensembles.
Figure 1. Median absolute change in temperature (°C) and percent change in precipitation and aridity index from 2010 to 2090 for Coupled Model Intercomparison Project Phase 5 (CMIP5) Representative Concentration Pathway (RCP) scenarios 4.5 and 8.5. Note the inverted color scale for temperature. There were no white or gray cells in the temperature plot when expressed in terms of percent change (not shown).

Several spatial trends in the maize yield response to climate change were robust across ensembles in spite of cross-ensemble differences in the emission scenario, the model generation (CMIP3 vs. CMIP5), and the ensemble size and type (within- or between-model). All ensembles projected that robust negative effects (orange) will be more widespread than robust positive effects (green). Regionally, all projected the most widespread yield losses in the Sahel region and southern Africa below the Democratic Republic of the Congo and Tanzania. Sub-regionally, all projected yield increases in the Ethiopian highlands, at the southern tip of South Africa, and (tentatively, as many grid cells in this region produce little to no maize and were therefore not modeled) in the Horn of Africa. No ensemble projected a substantial median yield change within or surrounding the tropical rainforests of Central and West Africa. As expected, the crop yield response increased over time, and the fraction of cells characterized by “no substantial effect” or “substantial but unknown effects” (e.g., gray or white cells) decreased.

The magnitude of the crop yield responses reflects the emission scenario, such that 2090 results for the SRES A1B between-model ensemble were bounded by those of RCP4.5 (lower emissions, smaller response) and RCP8.5 (higher emissions, larger response). This trend was preserved across the CMIP3 and CMIP5 between-model ensembles, suggesting that the scenario ultimately had a larger influence on results than the choice of model generation.
The ECHAM5 ensemble was relatively pessimistic and the NCAR ensemble was relatively optimistic. As expected, both of these within-model ensembles showed lower within-ensemble disagreement (white) than the three between-model ensembles. In the ECHAM5 ensemble, white persisted to the greatest extent in Tanzania, Kenya, and Sudan; in the NCAR model, white persisted in the Sahel region and in Southern Africa. It is likely that internal variability is responsible for much of the disagreement in these regions across the multi-model CMIP3 ensemble.

Comparison of Figures 1 and 2 reveals general agreement between projected changes in crop yields and projected changes in AI over the next century: yields decreased in regions that are expected to become drier (e.g., Southern Africa, the western Sahel region); yields increased in the arid/semi-arid Horn of Africa.
which is projected to become more humid; yields and aridity index (AI) both changed least in the tropical rainforest, at least in the RCP4.5 scenario; and the spatial distribution of white (uncertain) regions is similar for both AI and yields. The relative insensitivity of AI to climate change in the tropical rainforest results from high precipitation in this region.

Overall, maize yield trends aligned more closely with trends in AI (a composite index that describes the combined effect of changes in temperature and precipitation) than with trends in temperature or precipitation alone (the raw inputs to AquaCrop), illustrating the potential value of aridity as an index for evaluating maize yield changes with climate. Further, it should be noted that increases in aridity imply that irrigation would become less effective in these regions for the same amount of water use.

Figure 3 shows the 25th and 75th percentile (first and third quartile, or Q1 and Q3) percent change in maize yields from 2010 to 2090 in each grid cell across all ensemble members. The NCAR ensemble was not included because projections are not available for 2090. Each map represents an estimate of uncertainty based on these model runs, depicting a pessimistic and optimistic composite of future impacts under climate change. Note that an analysis of the ensemble members associated with each median and quartile in Figure 3 (not shown) suggests that no single ensemble member dominated results in any region of any subplot.

Spatial trends that are robust (green and orange) in Figure 2 also appear here in both the Q1 and Q3 maps for each ensemble. Thus, the direction of the crop yield response in these regions is not only robust across two thirds of the ensemble members but is also robust within the interquartile range. The Q1 map shows regions of substantial crop losses (greater than 25%) near the subtropical Sahel edge region extending across a line of latitude spanning West, Central, and East Africa in all model ensembles, as well as in parts of Southern Africa. Even in the Q3 map, substantial crop losses are projected for parts of Southern Africa in the CMIP5 RCP8.5 scenario. Comparison of the RCP 4.5 and 8.5 inner quartile maps illustrates the potential impact of emissions reductions on crop yields under optimistic and pessimistic conditions, discussed further below.

Unsurprisingly, for regions that are white in Figure 2, the projected direction of the crop yield response generally differs in Figure 3 across the Q1 and Q3 maps. Differences are especially pronounced in the Sahel region and northern East Africa. The light gray regions of Figures 2 and 3, which indicate small changes, robustly suggest that maize exhibits the lowest sensitivity to climate change in the rainforests of Central Africa.

The stacked bar charts in Figure 4 indicate the fraction of the harvested area in each geopolitical region (left) or aridity zone (right) for which the crop yield change from 2010 to 2050 was projected to fall into the color category defined in Table 1/Figure 2. The gray bar at the bottom of each map in the upper panel indicates the fraction of the total harvested area in SSA currently found within each region or zone. By region, maize production currently follows the order East > West > Southern > Central > Saharan. By zone, production follows the order humid > semi-arid > wet > arid (i.e., little production occurs in either the desert or rainforest).

We once again see good agreement across ensembles: All agreed that most of West and Central Africa will exhibit no substantial change. Three out of five agreed that East and Southern Africa, the first and third largest maize-producing regions in SSA, will experience the largest negative impacts. The two exceptions were the CMIP5 RCP4.5 and NCAR ensembles, which projected a robust negative response in East Africa but disagreed on the direction of change in Southern Africa (white). Notably, four out of five ensembles also projected that East, Southern, and Saharan Africa will see the largest robust positive impacts sub-regionally. The exception in this case was the NCAR ensemble, which projected larger yield increases than the other ensembles in both Central and Saharan Africa.

The largest maize-producing regions by far, the semi-arid and humid zones (~80% of total production) showed the largest robust negative signal. Negative robust effects were smallest in the “wet” zones. Within-ensemble disagreement (white) increased as zones become more arid. This is no surprise, since water-limited regions will be more sensitive than water-rich regions to the wide variation in projected precipitation changes that characterize GCMs (e.g., see Figure 1). Model sensitivity to temperature and precipitation changes is discussed in greater detail below.
Figure 3. Twenty-fifth percentile (first quartile, left column) and 75th percentile (third quartile, right column) percent change in yield in each cell across all ensemble members for each ensemble from 2010 to 2090. Note changes to the legend: because we are not concerned in this case with agreement across ensemble members, there are no white cells and gray cells are re-defined (for this figure only) as a change within 5%. NCAR ensemble results are not available for 2090.

Table 2 summarizes the potential effects of anthropogenic climate change in terms of the percent change relative to 2010 production in the total annual maize production in each region (where regions are defined as in Figure 4) that is projected to occur for three cases (1) 2090 RCP4.5 emissions, (2) 2090 RCP8.5 emissions, and (3) a switch from the CMIP5 RCP8.5 emissions pathway to that of RCP4.5 for 2090 for each of the 20 ensemble members found in both ensembles (i.e., mitigation). Results are also shown for all of SSA and eight of the nine countries that comprise 75% of the total current (2014) maize production in SSA according to FAO [2015], ranked from highest production (South Africa) to lowest production (Ghana). Malawi was not
Figure 4. (A) Sub-Saharan Africa divided into geopolitical regions (left) and aridity zones (right). The bar at the bottom of the map indicates the fraction of the total harvested area in sub-Saharan Africa found within each region or zone. (B) Median fraction of the harvested area in each region or zone that each ensemble project will fall into the categories described in Table 1 by 2050. Arid: $AI < 0.2$, semi-arid: $0.2 < AI < 0.5$, humid: $0.5 < AI < 0.75$, wet: $AI > 0.75$. $AI$ was based on the historical climate from 1997 to 2003 [Sheffield et al., 2006].

Included because it is represented in our model by a single cell. Given the coarse resolution of our model, which limits the ability of this and other crop models to reproduce historical maize yields at the country scale (Figure S5), we caution against over-interpretation of the country-scale results. However, they can provide a starting point for future analyses at smaller scales.

Figure S7 maps the absolute and percent change in production in every cell due to mitigation in case (3). Figure S8 shows Figure 4 results (percent change in the maize yield in each cell between 2010 and 2050 weighted by the fraction of harvested area and aggregated across all cells in the country) for the eight countries in Table 2.

Note that case (3) values do not equal the difference between columns (2) and (1) because (unlike for means) the difference in the quartiles of two data sets does not necessarily equal the quartiles of their differences. In
### Table 2. Current (2014) Production [Food and Agriculture Organization of the United Nations, 2015] and Percent Change in the Overall Maize Production by Region/Country Relative to 2010 Production for Three Cases: 2090 Levels for RCP4.5 and RCP8.5, and Emission Reductions From RCP8.5 Levels to RCP4.5 Levels in 2090

<table>
<thead>
<tr>
<th>Region/Country (Region Code)</th>
<th>Current Production (Megatonnes)</th>
<th>Case (1): Maize Change for 2090 Compared to 2010 Values for RCP4.5</th>
<th>Case (2): Maize Change for 2090 Compared to 2010 Values for RCP8.5</th>
<th>Case (3): Maize Gain for Mitigation From RCP8.5 to RCP4.5 Levels in 2090</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southern (S)</td>
<td>15 (−15, −7.2, 20)</td>
<td>(−27, −14, 4.4)</td>
<td>(−27, −14, 4.4)</td>
<td>(4.8, 16, 37)</td>
</tr>
<tr>
<td>East (E)</td>
<td>30 (−6.2, −1.9, 0.39)</td>
<td>(−13, −6.0, −0.22)</td>
<td>(−13, −6.0, −0.22)</td>
<td>(−3.0, 4.7, 8.6)</td>
</tr>
<tr>
<td>Saharan (Sh)</td>
<td>2.5 (−17, −8.7, 1.6)</td>
<td>(−28, −5.4, 12)</td>
<td>(−28, −5.4, 12)</td>
<td>(−14, −4.0, 8.9)</td>
</tr>
<tr>
<td>Central (C)</td>
<td>4.7 (−2.0, −0.87, 0.18)</td>
<td>(−5.6, −2.5, −0.74)</td>
<td>(−5.6, −2.5, −0.74)</td>
<td>(−0.29, 0.88, 3.8)</td>
</tr>
<tr>
<td>West (W)</td>
<td>18 (−4.2, −1.9, −0.11)</td>
<td>(−7.9, −4.1, −2.6)</td>
<td>(−7.9, −4.1, −2.6)</td>
<td>(−0.47, 2.5, 4.7)</td>
</tr>
<tr>
<td>South Africa (S)</td>
<td>14 (−18, −7.2, 20)</td>
<td>(−33, −23, 0.88)</td>
<td>(−33, −23, 0.88)</td>
<td>(8.5, 21, 38)</td>
</tr>
<tr>
<td>Nigeria (W)</td>
<td>11 (−2.9, −1.5, 0.82)</td>
<td>(−6.6, −2.4, −0.25)</td>
<td>(−6.6, −2.4, −0.25)</td>
<td>(−0.33, 1.8, 3.6)</td>
</tr>
<tr>
<td>Ethiopia (E)</td>
<td>7.2 (1.9, 5.8, 9.3)</td>
<td>(2.8, 13.21)</td>
<td>(2.8, 13.21)</td>
<td>(−12, −5.3, −0.98)</td>
</tr>
<tr>
<td>Tanzania (E)</td>
<td>6.7 (−6.5, −1.5, 4.2)</td>
<td>(−9.7, −2.4, 1.9)</td>
<td>(−9.7, −2.4, 1.9)</td>
<td>(−1.9, 0.12, 11)</td>
</tr>
<tr>
<td>Kenya (E)</td>
<td>3.5 (−12, 0.25, 17)</td>
<td>(−19, −2.2, 12)</td>
<td>(−19, −2.2, 12)</td>
<td>(−14, 0.24, 14)</td>
</tr>
<tr>
<td>Zambia (E)</td>
<td>3.4 (−15, −11, −2.0)</td>
<td>(−37, −22, −14)</td>
<td>(−37, −22, −14)</td>
<td>(4.6, 10, 23)</td>
</tr>
<tr>
<td>Uganda (E)</td>
<td>2.8 (−10, −1.3, 2.0)</td>
<td>(−13, −1.2, 6.4)</td>
<td>(−13, −1.2, 6.4)</td>
<td>(−5.8, −1.9, 8.7)</td>
</tr>
<tr>
<td>Ghana (W)</td>
<td>1.8 (−5.2, −1.1, 1.8)</td>
<td>(−12, −4.6, −0.47)</td>
<td>(−12, −4.6, −0.47)</td>
<td>(−0.97, 2.5, 6.7)</td>
</tr>
<tr>
<td>SSA, all</td>
<td>69 (−5.0, −2.9, −0.07)</td>
<td>(−12, −8.1, −2.6)</td>
<td>(−12, −8.1, −2.6)</td>
<td>(2.8, 5.5, 8.0)</td>
</tr>
</tbody>
</table>

RCP, Representative Concentration Pathway; SSA, sub-Saharan Africa. Lower and upper bounds represent the first and third quartile across ensemble members.

Climate change had detrimental effects on production, and mitigation had beneficial effects, in those regions and countries for which Figures 4 and S8 suggest robust agreement on the projected negative impacts of climate change on yields. We project particularly large regional benefits of mitigation (costs of climate change) in the high-producing regions of Southern and East Africa and particularly large national benefits (costs) in Zambia and South Africa. As expected, maize losses in RCP8.5 are generally larger than in RCP4.5. Two notable exceptions, Saharan Africa (<4% of current maize production in SSA) and Ethiopia (∼10%), suggest net positive median effects of climate change on production from RCP4.5 to RCP8.5. Ethiopia is a special case: due to larger yields in the cool highlands under warming (see below), it is the only country that experienced substantial positive changes in production under climate change.

Across all three cases, many regions and countries show disagreement on the direction of the production response to mitigation across the interquartile range. This disagreement arises from the regions of within-ensemble disagreement that are indicated in white and gray in Figure 2, which are included in the production calculation. Nationally, the potential benefits of mitigation are particularly uncertain for Kenya because all but one of the corresponding Figure 2 cells (for the CMIP5 RCP4.5 and RCP8.5 ensembles) are white.

### 3.3. Relative Influence of Climate Drivers on Yield Projections

Figure 5 presents the relative impacts of temperature and precipitation on projected changes in crop yields due to climate change from 2010 to 2090. The cells chosen for Figure 5 exhibited a substantial median yield change (i.e., are orange, green, or white) across at least three of the four ensembles for which 2090
Figure 5. Percent change in temperature (T), precipitation (P), and aridity index (AI) between 2010 and 2090 and their separate effects (\(\Delta T\) or \(\Delta P\), representing scenarios in which only \(T\) or \(P\) was varied from 2010 to 2090) and combined effects (\(\Delta T + \Delta P\), representing our realistic ‘base case’ results) on the crop response in eight representative grid cells from regions that show consistent trends across ensembles. Lower left (orange box): percent change in the total maize production in sub-Saharan Africa in the three scenarios. Results are for Coupled Model Intercomparison Project Phase 5 Representative Concentration Pathway 8.5 (spread represents 20 ensemble members). Each cell’s color matches its color in Figure 2. Boxplots represent the minimum, Q1, median, Q3, and maximum values, excluding outliers. Outliers [values greater than Q3 + 2.7\(\sigma\) (Q3 – Q1) or less than Q1 – 2.7\(\sigma\) (Q3 – Q1), where \(\sigma\) is the standard deviation] were omitted for visual clarity.

Projections are available (Figure 2). In addition, all cells were located within a sub-region that broadly exhibited the same response. To ensure that cells were representative of their sub-region, we corroborated all trends reported here for two adjacent cells with like colors. Figure S9 (described below) further explores AquaCrop’s sensitivity to temperature and precipitation in every cell in SSA, but only with respect to the median response across all ensemble members.

Boxplots on the left-hand side of each subplot in Figure 5 present the percent change in temperature, precipitation, and aridity index in each cell. As expected, temperatures increased at all locations, and variation in changes in the aridity index closely mapped variation in changes in precipitation. Variation in projected precipitation changes was large (ranging from positive to negative in all tested cells), but the median percent change in precipitation was generally smaller than that of temperature; in agreement with Schlenker and Lobell [2010], Figure S9 (upper left panel) reveals that the median temperature change exceeded the precipitation change in all but a few sub-regions (e.g., the Horn of Africa, the Sahara, and the southwest coast).

Boxplots on the right-hand side of each subplot show the percent change in crop yields when both temperature and precipitation were varied from the 2010 baseline (i.e., the base case reported previously). Results are also shown for two “partial derivative” scenarios in which only temperature or precipitation was varied. This approach allowed us to compare the relative effects of temperature and precipitation changes on crop yield changes from 2010 to 2090. The subplot in the bottom left, “SSA (all),” shows the percent change in the total production in SSA for the three scenarios. Figure S9 (upper right panel) reveals which climate driver (temperature, precipitation, or both) had the largest influence on the median projected crop yield change in every cell.

We found that yields increased in the Ethiopian highlands and the southern tip of Africa (two of the coldest sub-regions of Africa) due to increasing temperatures, and that yields increased in the Horn of Africa due
to increased precipitation. In the white (uncertain) cells, we found that temperature and precipitation had opposite effects on yields. In contrast, in the cells that showed negative effects, temperature and precipitation both had negative effects. Trends in aridity index generally mapped trends in yields more closely than trends in precipitation or temperature.

With respect to the relative influence of the two raw inputs, temperature and precipitation, Schlenker and Lobell [2010] and Challinor et al. [2014] previously found that crop production in Africa is more sensitive overall to temperature than to precipitation. The inset in Figure 5 (bottom left subplot) reveals that this is true in our model as well, but only with respect to the median change in the total production in SSA. The dominance of the temperature effect over all of SSA is driven by regional variation in the directionality of the yield response to temperature and precipitation, as shown in the lower panel of Figure 5: Temperature changes had negative effects on yields in all but two regions, the Ethiopian highlands and the southern tip. In contrast, precipitation changes had broadly negative effects in the southern and western regions of SSA and positive effects in the northeast; thus, much of the precipitation effect was canceled out when summed over all of SSA.

Overall, precipitation contributed most of the uncertainty to model projections, and, sub-regionally, either driver (or both) could control the crop response. Indeed, Figure 5 reveals that, for a given cell, the influence of precipitation on the yield change often exceeded that of temperature, even when the magnitude of the precipitation change in the cell was smaller than that of the temperature change. In a comparison of seven crop models, Eitzinger et al. [2013] found AquaCrop to be relatively insensitive to temperature changes and relatively sensitive to precipitation changes. This is unsurprising, since AquaCrop is also a “water-driven” crop model, as opposed to a “carbon-” or “radiation-driven” crop model [Todorovic et al., 2009]. It is therefore possible that our chosen crop model overestimates the influence of precipitation on future changes in crop yields.

In general, the spatial trends in the crop yield response projected in this work were robust to climate model uncertainties but are not robust to the choice of crop model. Figure S10 compares our results to those of three gridded crop models from the AgMIP ensemble [Rosenzweig et al., 2014]. Unsurprisingly, we find some substantial differences in the median crop yield response across the continent in the four models. Recall, however, that AquaCrop outperformed the AgMIP models for SSA at 2° × 2° resolution in a comparison of projections to historical estimates (Figure S5). A comparison of our projections to those of several recent works (below) suggests that, in spite of large and persistent uncertainties in climate change assessments for agriculture, our findings are in broad agreement with the existing literature.

4. Discussion

We used a process-based crop model (FAO AquaCrop-OS) and five “within-” and “between-model” GCM ensembles representing three emission scenarios as well as CMIP3 and CMIP5 to project climate-driven changes in SSA maize yields at 2° × 2° resolution in 2010, 2030, 2050, and 2090. We identified regions in which projections were robust across all five ensembles as well as regions where disagreement persists out to 2090 due to internal variability, scenario uncertainty, or model uncertainty. We grouped results by geopolitical region and aridity zone, explored the regional/national benefits of mitigation from an RCP8.5 to RCP4.5 (CMIP5) emission scenario, and explained regional and sub-regional trends in the crop yield response in terms of the underlying changes in temperature, precipitation, and aridity.

Maize yield projections were remarkably similar across all five ensembles. Notably, differences in the median response between the CMIP3 and CMIP5 ensembles were smaller overall in 2090 than the uncertainties introduced by the choice of emission scenario (RCP4.5 < SRES A1B < RCP8.5). In qualitative agreement with several recent cross-study comparisons, we project that climate change will have net negative impacts on maize production in SSA [Knox et al., 2012; Challinor et al., 2014; Rosenzweig et al., 2014]. We project the following median percent change in the total maize production in each region between 2010 and 2090 across the CMIP5 RCP4.5 and RCP8.5 (lowest and highest emission) scenarios. Southern Africa: −14 to −7.2%, Saharan Africa: −8.7 to −5.4%, East Africa: −6.0 to −1.9%, West Africa: −4.1 to −1.9%, Central Africa: −2.5 to −0.87%, and SSA (total): −8.1 to −2.9%. See Table 2 for further information on the impact of climate change and the potential benefits of mitigation of greenhouse gas emissions.
Our results broadly support previous findings that yield losses will be smallest in Central Africa [Lobell et al., 2008; Thornton et al., 2011; Waha et al., 2013; Rippke et al., 2016], relatively large in the Sahel region [Thornton et al., 2011; Waha et al., 2013; Rosenzweig et al., 2014; Rippke et al., 2016], and largest of all in the region south of the Democratic Republic of the Congo and Tanzania [Lobell et al., 2008; Schlenker and Lobell, 2010; Waha et al., 2013; Rippke et al., 2016]. South Africa, the largest maize producer in SSA (Table 2), could experience particularly large losses.

We project that increased rainfall will increase yields in the Horn of Africa and increased temperatures will increase yields in the Ethiopian highlands and at the southern tip of the continent. Waha et al. (2013) similarly projected crop yield increases in the Ethiopian highlands and at the southern tip. Thornton et al. [2011] placed the largest positive impacts (in terms of growing season length) in Kenya, and Rosenzweig et al. [2014] placed increases in the Horn of Africa; in any case, all four studies support the conclusion that crop production in sub-regions of northern East Africa is likely to benefit from climate change.

Our results show some important differences with previous work; others have projected larger negative impacts in West Africa than in Southern Africa [Thornton et al., 2011; Rosenzweig et al., 2014], and most studies find that negative impacts will be larger — though not dramatically so — in West Africa than in East Africa [Lobell et al., 2008; Thornton et al., 2011; Waha et al., 2013]. Differences arise in part because there is no consensus across studies on which countries belong in which region [Lobell et al., 2008; Thornton et al., 2011]. The magnitude of cross-study disagreement and the importance of crop model uncertainty are discussed further below.

A major aim of uncertainty analyses such as the one presented here is to guide future efforts to reduce risk and manage uncertainty. For all ensembles, disagreement on the sign of change across all ensemble members (white cells) was highest in the arid regions, followed by the high-producing semi-arid regions. White cell coverage (within-ensemble disagreement) in Figure 2 was only somewhat smaller in the within-model ensembles than the between-model ensembles, suggesting that internal variability is a major contributor to total climate model uncertainties. In general, internal variability led to substantial disagreement in the within-model ensembles even out to 2090 and explained most of the spatial distribution of uncertainty in current projections of maize yields under climate change in Africa. These findings agree with previous findings on internal variability as a major source of climate model uncertainty [Deser et al., 2012] and the inverse relationship between internal variability and model scale [Hawkins and Sutton, 2011]. Common modes of variability that contribute to internal variability in some regions in Africa include multi-year physical phenomena like ENSO (El Niño-Southern Oscillation), AMO (Atlantic Multi-decadal Oscillation), and the IOD (Indian Ocean Dipole) [Hoerling et al., 2006; Polo et al., 2008, 2011]. Internal variability is theoretically irreducible at long time scales.

Across much of the continent, most members of a given GCM ensemble agreed on the direction of change in the crop yield response. However, disagreement across the interquartile range in the Sahel region and the northern half of East Africa was quite large (Figure 3). The median and upper quartile estimates in Table 2 indicate substantial potential benefits of mitigation (substantial negative impacts of climate change). However, within-ensemble disagreement as reflected by the lower quartile projections also allows for the risk of costly rather than beneficial impacts of mitigation (beneficial rather than detrimental impacts of climate change) in most regions and countries. In our analysis, uncertainty resulted primarily from uncertainty regarding the direction of change in precipitation — a reflection of AquaCrop’s high sensitivity to precipitation and low sensitivity to temperature relative to other crop models and statistical analyses [Schlenker and Lobell, 2010; Eitzinger et al., 2013].

Differential sensitivities to temperature and precipitation are just one example of the ways in which the choice of crop model can have a large — and potentially dominant — impact on uncertainty in projected crop yield changes [Asseng et al., 2013; Wada et al., 2013; Rosenzweig et al., 2014]. By necessity, process-based gridded crop models employ many empirical approximations, ignore many factors that influence crop growth (e.g., pests, weeds, floods), and suffer from a lack of data for calibration and validation of sub-regional crop yield projections or parameterization based on regional management practices [Lobell et al., 2008; Müller et al., 2011; Knox et al., 2012; Rosenzweig et al., 2014]. Potential limitations of our crop model are discussed in the supporting methods (Section S5). For this study, which focuses exclusively on the sources and impacts of climate model uncertainty, support for FAO AquaCrop comes from its ability to
capture the two most important limitations on maize growth in SSA, water stress and nutrient limitations [Reynolds et al., 2015]; it performed well against three previous gridded crop models [Rosenzweig et al., 2014] in a comparison to historical data (Figure S5); and projections were in general agreement with previous work. Nonetheless, crop model uncertainty is an important target for future work.

We did not consider the uncertainties that would arise from a different choice of bias correction method or reference evapotranspiration equation, which can be substantial [Droogers and Allen, 2002; Ho et al., 2012]. In particular, our bias correction method only allows for mean changes in temperature and precipitation. Future work will consider changes in variability, which will allow us to explore seasonal trends. Notably, the magnitude of uncertainty in climate model projections is also a function of the ensemble size, with larger ensembles exhibiting lower variability in projections [Tebaldi and Knutti, 2007; Hawkins and Sutton, 2009; Knox et al., 2012].

In this work, we use simple "rules of agreement" (i.e., on the direction of change in the yield response) to extract robust trends from highly uncertain projections. Additional means of reducing uncertainty include (1) removing models from an ensemble of crop or climate models if they poorly describe critically important physical processes and (2) giving model projections with high skill for the current climate according to some chosen metrics of importance a larger weight in the analysis [Tebaldi and Knutti, 2007]. For example, Cook and Vizy [2006] demonstrated that only 10 out of 18 tested CMIP3 models exhibit a precipitation maximum during the West African monsoon season that falls over land rather than the ocean; these 10 models may be the best predictors of climate change impacts on maize yields in West Africa from CMIP3. Uncertainty can also be reduced via the development of smaller-scale assessments parameterized and validated with local or regional data on management and production [Challinor et al., 2014].

No matter the approach, uncertainty in projections of the far-future and sub-regional impacts of climate change on crop production is likely to remain high in the near term. Our work demonstrates that parts of Africa are subject to risks of large crop losses under climate change, especially at the upper end of the quartile range. Thus, our work identifies the need for risk management strategies that are adaptive and robust to uncertainty [Challinor et al., 2007; Hawkins and Sutton, 2009]. Those regions for which relatively small climate change impacts are projected, as discussed above, can rely on historical observations to advance "climate-proofing" to variability. For regions that could experience large impacts, Rippeke et al. [2016] suggest an incremental, multi-decadal, and multifaceted adaptation approach that combines near-term farm-scale changes, long-term geopolitical shifts, and real-time monitoring in order to create "a flexible enabling environment for self-directed change."

Farm-scale changes might include crop diversification and shifts in planting dates, cultivars, nutrient and pest management, and irrigation; large-scale institutional changes might include insurance programs, investment in research and agricultural extension, substitution of maize with sorghum or millet, and policy and infrastructure changes to support improved food transport, processing, and storage [Challinor et al., 2007; Lobell et al., 2008; Müller et al., 2011; Waha et al., 2013; Rippeke et al., 2016; van Ittersum et al., 2016]. Regions with robust yield losses also exhibited robust increases in aridity (i.e., decreasing AI), suggesting a promising role for irrigation in future adaptation strategies. Adaptation strategies aimed at increasing yields (as opposed to the expansion of harvested area) have particular promise in southern East Africa (e.g., Zambia and Zimbabwe), where projected climate change impacts on maize yields are especially large and current yield gaps (i.e., the gap between actual yields and potential yields in the absence of biophysical constraints such as nutrient and water limitations) are especially high [Mueller et al., 2012; van Ittersum et al., 2016].

Acknowledgments
We thank Brent Boehlert for providing the climate data and Andrzej Strzepek for assisting in data acquisition. All input data used in this analysis are publicly available, as cited. FAO AquaCrop-OS is open-source and has been extensively documented elsewhere (aquacropos.com). We acknowledge the CMIP3 and CMIP5 modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP’s Working Group on Coupled Modelling (WGCM) for their roles in making available the WCRP CMIP3 and CMIP5 multi-model datasets. For CMIP the U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. Funding for this research was provided by the Abdul Latif Jameel World Water and Food Security Lab (J-WAFS) at MIT.

References


Supporting Information for

Climate model uncertainty in impact assessments for agriculture: A multi-ensemble case study on maize in sub-Saharan Africa

Amy Dale\textsuperscript{1,2}, Charles Fant\textsuperscript{1}, Kenneth Strzepek\textsuperscript{1}, Megan Lickley\textsuperscript{2}, and Susan Solomon\textsuperscript{1,2}

\textsuperscript{1}Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, 77 Massachusetts Avenue, E19-411, Cambridge, MA 02139, USA. \textsuperscript{2}Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, 54-918, Cambridge, MA 02139, USA.

Corresponding author: Amy Dale (amydale@mit.edu)
Supporting Methods

This supplement provides details regarding the treatment of crop and soil parameter inputs, fertilizer and irrigation, and model calibration and benchmarking in AquaCrop-OS. Our treatment of the four daily climate time series inputs (minimum and maximum daily temperature, daily precipitation, and daily reference evapotranspiration) is described in the main text. Only methods specific to this work are presented here; we refer the reader to the FAO AquaCrop manuals for details regarding AquaCrop equations as well as definitions of common crop modeling parameters such as permanent wilting point, curve number, and water productivity [Foster, 2016; Raes et al., 2012].

S1. Crop parameters

All AquaCrop-OS v5.0a parameters describing (e.g.) crop growth and its sensitivity to temperature and water stresses were kept at their default values for maize. Planting and harvest dates for each grid cell were varied annually according to climate-driven rules based on [Waha et al., 2012].

Planting dates were either controlled by precipitation or by temperature according to whether or not a frost event was likely to occur during the year. In any given year, the crop was considered to be at risk of frost if the monthly average temperature ($T_{\text{month}}$) dropped below 10°C for at least one month of the year. If so, the planting month was taken to be the first month for which $T_{\text{month}}$ exceeded 14°C, and the planting day was taken to be the first day of this month for which the daily average temperature ($T_{\text{day}}$) exceeded 14°C. Note that the model was not sensitive to the choice of temperature threshold over the range considered, 10-14°C (results not shown).

If $T_{\text{month}}$ exceeded 10°C for all months, which was the case for the vast majority of grid cells in Africa as shown by [Waha et al., 2012], then the planting date was taken to be the first day after the start of the wet season for which the daily precipitation ($P_{\text{day}}$ in mm/d) exceeded 0.1 mm/d. The wet season was defined as the four-month time window with the highest sum ($AI_{\text{month}} = P_{\text{month}} / ET_{0,\text{month}}$), where $AI_{\text{month}}$ is the (dimensionless) monthly average aridity index and $P_{\text{month}}$ and $ET_{0,\text{month}}$ are the monthly average precipitation and reference evapotranspiration in mm/d, respectively. Note that we assumed that the wet season in any given year could start as early as January or as late as December, proceeding in the latter case through March of the subsequent year. In this case, since months that fell within a previous year’s wet season were also part of its growing season, they were not potential candidates for planting in the subsequent year.

This algorithm simplifies the planting rules proposed by [Waha et al., 2012]. Notably, by using averages from the current year in place of exponential weighted moving averages that account for past years, we implicitly assume that farmers have perfect knowledge of the current and future weather patterns for the given year and therefore do not need to base their planting decisions on expectations from previous years. In spite of this simplification, the planting months chosen with our method agree within 1-2 months with upscaled estimates of historical planting dates from the FAO/IIASA GAEZ Data Portal [FAO, 2016] (Figure S1). Substantial (4-6 month) differences are observed in the region dominated by rainforest (red). However, these
differences are deceptive, since there is very little maize production in this region and the growing season in the surrounding regions is very long (210-270 days) [Vrielings, 2013]. [Waha et al., 2012] additionally note that crop models are not sensitive to the choice of planting date near the equator.

The latest possible harvest date was set to be thirty days after the default number of days that maize requires to reach physiological maturity (132 days). 152 days is a low estimate of the maximum growing season length in Africa for some regions [Vrielings, 2013]. However, we found that the model was not sensitive to extensions of the growing season.

S2. Soil parameters

All AquaCrop-OS v5.0a soil parameters were kept at their default values, except where noted. The soil was divided into two layers with different compositions and hydraulic properties: topsoil (0 to 30 cm) and subsoil (30 cm to depth). For the purpose of maintaining an accurate soil water balance, layers were further subdivided into 18 compartments that expanded in size from 1 cm at the surface to 30 cm at depth. The total simulated soil depth was 2.6 m, where 2.3 m was the (default) maximum root depth for maize.

We simulated the soil water balance during dormant seasons as well as growing seasons in order to capture the ability of off-season climate to positively or negatively impact crop germination and growth in subsequent growing seasons. The initial water content of each soil layer was assumed to be 50% of the total available water (TAW, or the difference between the soil water content at field capacity and the water content at the wilting point). Because we simulated the soil water balance for five simulation years before collecting results (‘model spin-up’), our maize yield projections were not sensitive to assumptions about the initial soil water content.

We downloaded the following soil parameters from the Regridded Harmonized Soil Database v1.2 [Wieder et al., 2014]: topsoil and subsoil organic carbon (% weight), bulk density (g/cm³), and the fractions of sand, clay, and silt (% weight). Data were upscaled from 0.05° × 0.05° to 2° × 2° by averaging across all the subcells in each 2° × 2° cell.

We estimated soil hydraulic properties using the following pedotransfer functions [Mohamed and Ali, 2006; Nemes et al., 2005]:

\[ sat = -0.8667*Cl - 1.426*Sa - 84.2817*BD - 0.0151*(Si^2) + 0.0012*(Sa^2) \]
\[ + 7.9188*ln(Si) + 112.0333*ln(BD) - 0.0064*Cl*Si - 0.2835*Cl*BD \]
\[ - 0.0608*Si*Sa + 0.177*Si*OM + 266.768 \] (S1)

\[ FC = 0.0023*Si^2 - 8.491*BD^2 - 3.2498*OM^2 + 153.59021/Cl - 101.21431/Si \]
\[ - 9.02181/Sa + 8.52011/OM + 20.9002*ln(OM) + 0.3555*Cl*BD \]
\[ - 0.2388*Cl*OM + 10.1357*BD*OM + 16.4788 \] (S2)

\[ WP = -0.7409*Si + 0.0126*Si^2 - 7.4396*BD^2 - 2.8807*OM^2 + 136.151/Cl \]
\[ - 98.33231/Si - 23.99671/Sa + 9.03681/OM + 20.7999*ln(OM) \]
\[ + 0.4598*Cl*BD - 0.2579*Cl*OM + 9.1905*BD*OM + 9.8444 \] (S3)
\[
\ln(HC) = 7.75 + 0.0352*Si + 0.93*Top - 0.967*BD^2 - 0.000484*Cl^2 \\
- 0.000322*Si^2 + 0.001/\text{Si} - 0.0748/\text{OM} - 0.643*\ln(Si) - 0.01398*BD*Cl \\
- 0.1673*BD*OM + 0.02986*Top*Cl - 0.03305*Top*Si
\]  \hspace{1cm} (S4)

where

\begin{itemize}
  \item \textit{Sat} = water content at saturation (m$^3$/m$^3$)
  \item \textit{FC} = water content at field capacity (m$^3$/m$^3$)
  \item \textit{WP} = water content at permanent wilting point (m$^3$/m$^3$)
  \item \textit{HC} = hydraulic conductivity (cm/d)
  \item \textit{BD} = bulk density (kg/dm$^3$)
  \item \textit{Sa} = sand content (% weight)
  \item \textit{Si} = silt content (% weight)
  \item \textit{Cl} = clay content (% weight)
  \item \textit{OM} = organic matter content (% weight)
  \item \textit{Top} = dummy variable indicating whether (\textit{Top} = 1) or not (\textit{Top} = 0) the predictor variables represent the topsoil
\end{itemize}

The organic matter content was calculated from the total organic carbon content of the soil assuming the standard relationship, \( OM = 1.72*OC \). \textit{Sat}, \textit{FC}, and \textit{WP} were not allowed to fall below 0 or exceed 1.

The curve number associated with Antecedent Moisture Class (AMC) II (i.e., halfway between \textit{WP} and \textit{FC}) was estimated from the hydraulic conductivity of the topsoil according to the procedure outlined by [\textit{U.S. DOI}, 1993] (see Figure 2.10 of that work). AquaCrop adjusts this curve number in response to changes in AMC, or topsoil moisture.

\section*{S3. Fertilizer Application}

The total nitrogen, phosphorus, and potassium (NPK) fertilizer applied to maize in each grid cell in kilograms per hectare of harvested maize area was downloaded from EarthStat (www.earthstat.org) and upscaled from 0.05° x 0.05° to 2° x 2° by taking the weighted average of all subcell values, where the weighting factor was the fraction of the cell’s harvested maize area in hectares (also from EarthStat) represented by the given subcell [\textit{Monfreda et al.}, 2008; \textit{Mueller et al.}, 2012].

In addition, EarthStat estimates of the attainable yield \( (Y_{att}) \) and the yield gap \( (Y_{gap} = Y_{att} - Y_{act}) \), where \( Y_{act} \) is the yield that is actually achieved) were used to calculate the percent of the attainable yield that was achieved in each cell in previous (census-based) baseline estimates [\textit{Monfreda et al.}, 2008]. The attainable yield is the highest yield that can be achieved in a grid cell as approximated from the highest yields observed in regions with similar climates. Attainable yields are related to potential yield, the highest yield that can be achieved in a cell given biophysical constraints [\textit{Mueller et al.}, 2012].

\[%Y_{att} = \frac{(Y_{act}/Y_{att})}{100}\% = \frac{[Y_{att} - Y_{gap}]/Y_{att}}{100}\% \hspace{1cm} (S5)\%

\%Y_{att} \text{ was upscaled from 0.05° x 0.05° to 2° x 2° by taking the weighted average of the sub-cell values with respect to the harvested area, as described above.}
We then used linear regression to establish the relationship between the NPK application rate and \( %Y_{\text{att}} \) (i.e., the ability of fertilizer to close the yield gap). Since yield gaps are primarily caused by both nutrient stress and water stress, we isolated the influence of nutrient stress from that of water stress (to the extent possible) by excluding from the regression cells for which less than 10% of the maize harvested area in the cell was equipped for irrigation (see section S4). Exclusion based on irrigation rather than precipitation (also adopted by [Deryng et al., 2011]) has strengths and weaknesses: Irrigation is a stronger indicator than precipitation, since the ability of precipitation to reduce water stress depends to a large extent on (e.g.) on soil properties and aridity and also exhibits non-linear effects, with high-rainfall regions exhibiting disproportionately poor yields due to (e.g.) flooding and nutrient leaching as well as confounding factors (e.g., soil acidity, striga weed infestations, and poor terrain suitability) [Fischer et al., 2012; Khan et al., 2014; Reynolds et al., 2015; Stoorvogel et al., 1993]. However, only including cells that receive relatively high irrigation inputs is likely to bias projections of \( %Y_{\text{att}} \) upwards by favoring regions characterized by high-intensity agriculture that also employ more advanced farming equipment, pesticide, etc.

Cells were also excluded from the regression analysis if the dependent variable (\( %Y_{\text{att}} \)) was unavailable for 90% of subcells. In contrast, in order to ensure wide coverage of the continent, it was necessary to set missing subcell fertilizer and irrigation values to zero. This assumption had little to no effect on the upscaled grid values used for the regression because the subcells in question corresponded with regions of little to no harvested area.

The results of the regression are shown in Figure S2. The regression equation (\( R^2 = 0.41 \)) is:

\[
%Y_{\text{att}} = 0.229 \times (NPK) + 18.5
\]

(S6)

where \( NPK \) is the total nitrogen, phosphorus, and potassium fertilizer application rate in kg/ha.

We do not use Eq. S6 in our model to directly estimate reductions in yield due to fertilizer application. Rather, as described in Section S6, we use this equation during model calibration in order to determine the values of two influential crop growth parameters in each cell, maximum canopy cover and water productivity.

**S4. Irrigation**

The fraction of the total land area in each grid cell that is actually irrigated is the product of the fraction equipped for irrigation and the fraction of the equipped fraction that is actually irrigated (which may be smaller than the equipped area due to, e.g., lack of water or equipment damage), as defined and estimated at 0.05° x 0.05° in FAO AQUASTAT’s Global Map of Irrigation Areas v5. Note that this value includes areas equipped for full/partial control irrigation, spate irrigation, and equipped lowlands but does not include water management more broadly, including cultivated wetlands, inland valley bottoms, and non-equipped flood recession cropping [Siebert et al., 2013; Siebert et al., 2005]. In order to convert the irrigated fraction of the total area in the grid cell to the irrigated fraction of the harvested maize area, we divided the former (irrigated hectares / total hectares) by the harvested maize area fraction of the cell (harvested maize hectares / total hectares) as reported by EarthStat [Monfreda et al., 2008]. In so doing, we assumed that maize has first priority for irrigation. This is a reasonable first estimate, since
maize is not only the most widely produced crop in Africa with respect to total harvested area but is also relatively water-intensive and drought-sensitive; overall, cereals comprise the largest fraction of irrigated crop area in all regions of Africa except the Gulf of Guinea, where irrigated area is dominated by vegetables [Frenken, 2005].

The fraction of the harvested maize area equipped for irrigation was capped at one. Estimates were then upscaled to 2° x 2° using area-weighted averaging as described for fertilizer.

Crop yields reported in the Results (main text), $Y_i$, were calculated as follows,

$$Y_i = f_{irr,i} * Y_{irr,i} + (1-f_{irr,i}) * Y_{noirr,i}$$

(S7)

where $f_{irr,i}$ is the fraction of the harvested maize area in cell $i$ that is irrigated and $Y_{irr,i}$ and $Y_{noirr,i}$ are the yields for “complete irrigation” and “no irrigation” scenarios, respectively. The “complete irrigation” scenario assumed no water stress. Specifically, we used the AquaCrop ‘net irrigation water requirement mode’ to keep the soil water content at field capacity during the entire simulation.

Taken together, the simplifying assumptions described here are likely to somewhat overestimate $f_{irr,i}$. Unfortunately, a “first-order assessment” of this sort is necessitated by the current lack of sub-national data on irrigation application rates by crop and fodder type. In spite of these limitations, we included irrigation because it substantially improved the model’s ability to project historical crop yields in major maize-producing regions during calibration. Assumptions regarding irrigation will be revisited in future work.

S5. Additional Assumptions

Since yields are zero by definition in regions with no maize production, they were also set to zero in our model for any cell in which 90% of subcells (at 0.05° x 0.05° resolution) had no harvested maize area. The potential maize yield impacts of sequential cropping and (especially) intercropping, which comprise a significant fraction of all sub-Saharan African maize production [Waha et al., 2013], were not considered. Although we only model a single growing season, the Guinea coast and the Horn of Africa have two possible growing seasons due to a bimodal rainfall distribution.

AquaCrop accounts for aeration stress when the maize root zone becomes water logged; however, it does not explicitly model flooding. This inability to describe flooding could be a problem for lowland valleys in the tropics [Fischer et al., 2012]. Notably, model projections were not substantially improved by the explicit inclusion of a groundwater table (results not shown); in light of model insensitivity to groundwater table effects on water logging and capillary rise, as well as the uncertainty surrounding current estimates of the depth to groundwater across sub-Saharan Africa [Fan et al., 2013], the groundwater table was omitted from the model.

Maize yields in sub-Saharan Africa are primarily limited by the abiotic constraints modeled in this work: nutrient limitations and sensitivity to drought [Reynolds et al., 2015]. Like other gridded crop models [Rosenzweig et al., 2014], AquaCrop does not account for biotic constraints such as insect pests and weeds. Potential for AquaCrop and similar models to overpredict maize
yields in low-yielding regions results not only from the omission of biotic factors in the model but the omission of several positively correlated constraints that reduce maize production in the tropics and sub-tropics, including striga infestations, native soil quality (e.g., pH), and terrain suitability [Fischer et al., 2012; Khan et al., 2014; Reynolds et al., 2015; Stoorvogel et al., 1993]. Systematic biases of this sort are reduced, but not eliminated, by model calibration.

### S6. Model calibration and benchmarking

We calibrated AquaCrop by minimizing the root-mean-square error (RMSE) between national year 2000 average crop yields estimated by AquaCrop for sub-Saharan Africa (i.e., excluding Algeria, Egypt, Libya, Morocco, and Tunisia) and national averages reported by FAOSTAT for the same year [FAO, 2015]. FAOSTAT data were selected for calibration because they are based on national surveys and are thus likely to be more accurate than aggregated sub-national estimates such as those from FAO/IIASA GAEZ [FAO, 2016]. The AquaCrop year 2000 estimates were the average of seven years of annual yields for a simulation driven by historical (Princeton) climate data from 1997-2003 [Sheffield et al., 2006]. We used five years of spin-up (climate years 1991-1996) to estimate the soil water content at the beginning of the growing season associated with the 1997 harvest, which was sown in either late 1996 or 1997 depending on the grid cell (see Section S1 on planting date rules).

AquaCrop national averages were the area-weighted estimates across all grid cells that fell within the boundaries of a given nation. Cells were simply apportioned to a nation if most of their land area fell within that nation; in other words, cells were not apportioned to neighbors based on the fraction of land area that fell within each nation. Figure S3 shows the relationship between cells and country boundaries.

Two global parameters were “tuned” during model calibration in order to minimize RMSE. As described shortly, both were shape parameters describing the relationship between fertilizer application and crop growth. We chose this method for two reasons: First, it is a reduced form of the recommended soil fertility-based calibration method suggested in the AquaCrop manual [Raes et al., 2012]. Second, it reflects the critical, even dominant, influence of soil fertility on maize yields in sub-Saharan Africa [Reynolds et al., 2015; Sanchez, 2002]. In light of the limited data available for calibration, we decided to distill the recommended site-scale and data-intensive AquaCrop calibration methods down to their most essential components and tune the model globally rather than cell-by-cell. This approach reduced the risk of model overparameterization, which would have resulted in a model with high accuracy during calibration but low predictive power.

Eq. S6 describes the relationship between the NPK fertilizer application rate (a proxy for soil fertility) and the percent of the attainable maize yield achieved, $%Y_{at}$, where the attainable yield is taken to be the yield that would be projected by AquaCrop in the absence of soil fertility constraints on crop growth. As recommended in the AquaCrop manual, $%Y_{at}$ is then translated into “soil fertility stress,” $SF$, a general descriptor of soil quality that ranges from zero (no stress) to one (complete stress). We assume a simple inverse relationship:

$$SF = 1 - \frac{Y_{at}}{100\%}$$  \hspace{1cm} (S8)
SF is related in turn to a percent reduction in several crop growth parameters. Each curve describing this reduction takes the general form

\[ K_{adj,x} = p_x SF^2 - SF*(p_x + 1), \quad -1 \leq p_x \leq 1 \]  

(S9)

where \( K_{adj,x} \) is the downward adjustment factor for crop growth parameter \( x \) (percent reduction in \( x = 100%*K_{adj,x} \)) and \( p_x \) is a fitting parameter specific to \( x \). Note that Eq. S9 describes a curve that passes through the points (0, 1) – no stress, no adjustment—and (1, 0) – complete stress, no crop growth—and is either concave down (\( p < 0 \)), linear, or concave up (\( p > 0 \)).

The calibration method in the AquaCrop manual applies curves of the form described by Eq. S9 to four crop parameters. We simplified by adjusting only the two most sensitive parameters, the maximum crop cover (CCx) and the crop water productivity (WP). Our global calibration parameters were the \( p_x \) values, \( p_{CCx} \) and \( p_{WP} \). In effect, varying these two parameters allowed us to alter the global sensitivity of the model to nutrient stress in order to find the closest agreement between model results and historical estimates. The adjustment curves resulting from the model calibration procedure are shown in Figure S4.

We used a simple linear search procedure to calibrate the model. In order to reduce the influence of single grid cells and the smallest maize producers on model fits, countries were only included in the calibration if they were represented by more than one grid cell and had a total harvest maize area above \( 10^4 \) hectares (\( n=27 \) countries). For completeness, however, the \( R^2 \) values reported below include all countries for which data are available, regardless of size (\( n=33 \) countries).

Figure S5 shows the agreement between our AquaCrop model projections and historical estimates. In addition to survey-based national estimates from FAOSTAT (used for model calibration), we also compared results to aggregated sub-national estimates from [FAO, 2016] and [Monfreda et al., 2008] (as reported at the country scale by [Mueller et al., 2012]).

Moreover, we compared AquaCrop’s performance with that of three other process-based gridded crop models from the 2012 AgMIP (Agricultural Model Intercomparison and Improvement Project) “Fast Track” project [Rosenzweig et al., 2014]: LPJmL, pDSSAT, and PEGASUS. In order to make a direct comparison between our model and the AgMIP models, we used AgMIP results for the “no CO2 effects” scenarios and weighted their results for the “no irrigation” and “complete irrigation” cases by our estimate of \( f_{irr} \), using Eq. S7. Importantly, we also upscaled AgMIP model results to 2° x 2° by simple (unweighted) averaging in order to make model scales comparable with our own; the AgMIP models are likely perform somewhat better at their native 0.5° x 0.5° resolution.

Overall agreement between observed historical national averages and aggregated AquaCrop estimates was not particularly high, with \( R^2 \) values falling between 0.2 and 0.4. This weak-to-moderate agreement is not unexpected given the many complex biotic and abiotic factors that control maize growth. Nonetheless, AquaCrop’s performance was noticeably better than that of two of the three AgMIP models, PEGASUS and pDSSAT, which performed quite poorly. \( R^2 \) values for LPJmL were comparable to those for AquaCrop when compared to FAOSTAT data and estimates by [Monfreda et al., 2008]; however, LPJmL exhibited higher bias overall and perform relatively poorly against the GAEZ data.
Figure S6 labels the top left plot in Figure S5 (AquaCrop vs. FAOSTAT) by country. AquaCrop’s over-predictions of Ethiopian maize yields and under-predictions of South African and Tanzanian maize yields are the most problematic in light of both the magnitude of the error and the relative importance of these countries to the total maize production in sub-Saharan Africa (e.g., see Table 2, main text). In general, it is clear that the current generation of gridded crop models have a limited ability to project historical country-scale crop yields, at least at 2° x 2° resolution. For two reasons, this sobering conclusion does not preclude the usefulness of these models in first assessments of potential climate impacts at the regional (multi-national) scale. First, relative changes in yields are more robust to errors than predictions of absolute maize yields [Rosenzweig et al., 2014]. Second, the most likely cause of poor agreement between historical and modeled maize yields is the simplification and omission of factors that influence crop growth besides nutrient limitations and climate. These factors will only affect estimates of the percent change in maize yields due to a changing climate—the overriding focus of this work—to the extent that they are themselves affected by climate change. To this end, gridded crop models designed for climate change impact assessment could benefit greatly from better scientific understanding of the relationships between climate and secondary but still substantial constraints on maize production in sub-Saharan Africa such as insect pests, weeds, and flooding. Model calibration is a valuable exercise in any case because it corrects to some degree for systematic biases in the baseline (historical) water demand: For example, if the model were to overestimate crop water demand by overestimating historical maize yields, it would be systematically biased upwards in its estimate of the negative impacts of future drought conditions.
### Table S1. Models, modeling groups, and SRES scenarios employed from the CMIP3 ensemble

<table>
<thead>
<tr>
<th>CMIP3 GCM</th>
<th>Country of Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bjerknes Centre for Climate Research, Bergen Climate Model 2.0</td>
<td>Norway</td>
</tr>
<tr>
<td>Center for Climate Modeling and Analysis, Coupled GCM 3.1</td>
<td>Canada</td>
</tr>
<tr>
<td>Center for Climate Modeling and Analysis, Coupled GCM 3.1.t63</td>
<td>Canada</td>
</tr>
<tr>
<td>Centre National de Recherches Météorologiques, Coupled Model 3</td>
<td>France</td>
</tr>
<tr>
<td>Commonwealth Scientific and Industrial Research Organization, Mk 3.0</td>
<td>Australia</td>
</tr>
<tr>
<td>Commonwealth Scientific and Industrial Research Organization, Mk 3.5</td>
<td>Australia</td>
</tr>
<tr>
<td>Geophysical Fluid Dynamics Laboratory, Climate Model 2.0</td>
<td>US</td>
</tr>
<tr>
<td>Geophysical Fluid Dynamics Laboratory, Climate Model 2.1</td>
<td>US</td>
</tr>
<tr>
<td>Goddard Institute for Space Studies, Atmospheric Ocean Model</td>
<td>US</td>
</tr>
<tr>
<td>Goddard Institute for Space Studies, ModelEH</td>
<td>US</td>
</tr>
<tr>
<td>Goddard Institute for Space Studies, ModelER</td>
<td>US</td>
</tr>
<tr>
<td>Institute of Atmospheric Physics, Climate Model System FGOALS G 1.0</td>
<td>China</td>
</tr>
<tr>
<td>Institute for Numerical Mathematics, Climate Model 3.0</td>
<td>Russia</td>
</tr>
<tr>
<td>Institut Pierre Simon Laplace, Climate Model 4</td>
<td>France</td>
</tr>
<tr>
<td>Model for Interdisciplinary Research on Climate 3.2, High Resolution</td>
<td>Japan</td>
</tr>
<tr>
<td>Model for Interdisciplinary Research on Climate 3.2, Medium Resolution</td>
<td>Japan</td>
</tr>
<tr>
<td>Max Planck Institute for Meteorology, European Center Hamburg Model 5</td>
<td>Germany</td>
</tr>
<tr>
<td>Meteorological Research Institute, Coupled General Circulation Model 2.3.2a</td>
<td>Japan</td>
</tr>
<tr>
<td>National Center for Atmospheric Research, Community Climate System Model 3.0</td>
<td>US</td>
</tr>
<tr>
<td>National Center for Atmospheric Research, Parallel Climate Model</td>
<td>US</td>
</tr>
<tr>
<td>UK Met Office, Hadley Center Climate Model 3</td>
<td>UK</td>
</tr>
<tr>
<td>UK Met Office, Hadley Center Global Environmental Model 1</td>
<td>UK</td>
</tr>
</tbody>
</table>
Table S2. Models, modeling groups, and emissions scenarios employed from the CMIP5 ensemble

<table>
<thead>
<tr>
<th>Modeling Center (or Group)</th>
<th>CMIP5 GCM</th>
<th>RCP 4.5</th>
<th>RCP 8.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia</td>
<td>ACCESS1-0</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>ACCESS1-3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Beijing Climate Center, China Meteorological Administration</td>
<td>bcc-csm1-1</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Canadian Centre for Climate Modelling and Analysis</td>
<td>CanESM2</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>National Center for Atmospheric Research</td>
<td>CCSM4</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence</td>
<td>CSIRO-Mk3-6-0</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>EC-EARTH consortium</td>
<td>EC-EARTH</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>The First Institute of Oceanography, SOA, China</td>
<td>FIO-ESM</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>NOAA Geophysical Fluid Dynamics Laboratory</td>
<td>GFDL-CM3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>GFDL-ESM2G</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>GFDL-ESM2M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>NASA Goddard Institute for Space Studies</td>
<td>GISS-E2-H</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>GISS-E2-H-CC</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>GISS-E2-R</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>GISS-E2-R-CC</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>National Institute of Meteorological Research/Korea Meteorological Administration</td>
<td>HadGEM2-AO</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)</td>
<td>HadGEM2-CC</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology</td>
<td>MIROC5</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies</td>
<td>MIROC-ESM</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>MIROC-ESM-CHEM</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology)</td>
<td>MPI-ESM-LR</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Norwegian Climate Centre</td>
<td>NorESM1-M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>NorESM1-ME</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
**Figure S1.** Comparison of planting months estimated by IIASA (left) to the average planting months predicted by our model over the calibration period, 1997-2003 (middle). IIASA data are not available for cells with little to no maize production. Right: Difference (in months) between IIASA planting months and those in our model.

**Figure S2.** Relationship between the total nitrogen, phosphorus, and potassium fertilizer application rate in kg/ha ($NPK$) and the percent of the attainable maize yield achieved in each cell ($%Y_{att}$). Results were used for model calibration.
Figure S3. Simple $2^\circ \times 2^\circ$ grid cell delineation for the calculation of national average maize yields

Figure S4. Relationship between soil fertility stress, $SF$, and adjustment factors ($K_{adj,x}$) for the maximum canopy cover ($x = CC_{max}$) and water productivity ($x = WP$) as determined by calibration. In Eq. S9, $p_{CCx} = -0.5$ and $p_{WP} = -0.4$. 
Figure S5. Results of the model calibration and comparison to results for global gridded crop models. Each data point represents a country and is sized by the country’s hectares of harvested maize area. Solid line: 1:1 line. Red dashed line: best fit line.
Figure S6. AquaCrop vs. FAOSTAT (top left subplot in Figure S5) with data points labeled by country. Labels were not included for countries with very little harvested maize area if they overlapped with the labels for countries with a larger harvested maize area. Solid line: 1:1 line. Red dashed line: best fit line.
Figure S7. 25th percentile (Q1), median, and 75th percentile (Q3) absolute change (upper panel) and percent change (lower panel) in 2090 maize production relative to 2010 production across all ensemble members (production = yield · harvested area) for mitigation from the CMIP5 RCP8.5 emission scenario to the RCP4.5 scenario. Dark grey cells were not modeled.
Figure S8. Median fraction of the harvested area in eight of the nine countries comprising 75% of the total current (2014) maize production in SSA that each ensemble projects will fall into the categories described in Table 1 (main text). Countries are ordered from highest (South Africa) to lowest (Ghana) current production [FAO, 2015]. Malawi was not included because it is modeled by a single cell.
Figure S9. *Upper panel:* Relative magnitude of the median percent change in temperature and precipitation ($\Delta T$ and $\Delta P$, left) and their influence on the percent change in yield ($\Delta Y$, right) from 2010 to 2090 for the CMIP5 RCP8.5 ensemble. *Left:* Colors indicate whether $\Delta T$ exceeds $\Delta P$ (red) or vice versa (blue) in each cell. *Right:* Colors indicate whether $\Delta Y$ for the base case scenario (“$\Delta T + \Delta P$”—see Figure 5 for a description of the scenarios) is closer to $\Delta Y$ for the “$\Delta T$ only” scenario (red) or $\Delta Y$ for the “$\Delta P$ only” scenario (blue), i.e., whether $\Delta T$ or $\Delta P$ (respectively) drives $\Delta Y$. If the differences between $\Delta Y$ for the “$\Delta T$ only” and the “$\Delta P$ only” scenarios and the “$\Delta T + \Delta P$” scenario both fall within ±20% of their average difference, both inputs are assumed to influence results (light grey). Dark grey cells were not modeled. *Lower panel:* Colors indicate whether the median yield change in each cell ($\Delta Y$) was positive (green), negative (orange), or zero (grey) in the “$\Delta T$ only” scenario (left) and the “$\Delta P$ only” scenario (right).
Figure S10. Median percent change in maize yields from 2010 to 2010, 2030, 2050, and 2090 for the AquaCrop model run with the 20-member CMIP5 RCP8.5 GCM ensemble and three global gridded crop models from the AgMIP FastTrack project run with a 5-member subset of the CMIP5 RCP8.5 ensemble (Rosenzweig et al. 2015). Note that the difference in ensemble size will affect the comparison. Color rules are described in Table 1 (main text).
Literature Cited


