Dear Editor,

Please find enclosed our revised manuscript. We have addressed the reviewers’ comments as outlined in the responses to the reviewers uploaded earlier. In some cases, restructuring of the manuscript and changes in results have meant that our text changes deviate somewhat from our earlier response (all changes are highlighted in the marked-up manuscript version).

Overall, we have restructured the paper and collected all methodical descriptions that had previously been scattered across sections 2, 3, and 4, now in one place in section 2. We have clarified the methods to better explain the differences between the three emulator approaches.

In the first submission, the emulators had been trained using only data from RCP8.5. After adding the significance test as suggested by the reviewers, and doing some more testing, we have decided to use all available data from all four RCPs to train the emulator. This has improved overall emulator performance and has led to extensive changes to the results in section 3.2 and 4. Results for the emulators trained only on RCP8.5 are still available in the Supplement, and the effect of using either the full or limited training data is discussed at several points in sections 3.2 and 4.

Moreover, we discovered that pDSSAT (one of the crop models) deviated from the simulation protocol which meant that we had to redo the analysis for this model. We have added descriptions of the differences between model settings of pDSSAT and the other models in section 2 and added text to the results sections where these differences in simulation setup lead to differences in the results.

Attached to this letter you will find a marked-up manuscript version showing all changes to the text. In addition to the text changes, we have also updated all figures. Figure updates include better colour scales, increased font sizes for the labels, and improved figure layouts to make better use of the available space.

We believe that the manuscript has improved substantially thanks to the reviews, and hope that it is now acceptable for publication in ESD. We look forward to your response, and remain at your disposal for any further questions.

For the authors,

Sebastian Ostberg, Jacob Schewe
Changes in crop yields and their variability at different levels of global warming

Sebastian Ostberg1,2, Jacob Schewe1, Katelin Childers1, and Katja Frieler1
1Potsdam Institute for Climate Impact Research, Telegrafenberg A31, 14473 Potsdam, Germany
2Geography Department, Humboldt-Universität zu Berlin, Berlin, Germany
Correspondence: Sebastian Ostberg (ostberg@pik-potsdam.de)

Abstract. An assessment of climate change impacts at different levels of global warming is crucial to inform the political discussion about mitigation targets, as well as for the economic evaluation of climate change impacts e.g. in economic models such as Integrated Assessment Models (IAMs) that internally use global mean temperature change as indicator of climate change. There is already a well-established framework for the scalability of regional temperature and precipitation changes with global mean temperature change ($\Delta$GMT). It is less clear to what extent more complex, biological or physiological impacts such as crop yield changes can also be described in terms of $\Delta$GMT; even though such impacts may often be more directly relevant for human livelihoods than changes in the physical climate. Here we show that crop yield projections can indeed be described in terms of $\Delta$GMT to a large extent, allowing for a fast interpolation of crop yield changes to emission scenarios not originally covered by climate and crop model projections. We use an ensemble of global gridded crop model simulations for the four major staple crops to show that the scenario dependence is a minor component of the overall variance of projected yield changes at different levels of $\Delta$GMT. In contrast, the variance is dominated by the spread across crop models. Varying CO$_2$ concentrations are shown to explain only a minor component of the remaining crop yield variability at different levels of global warming. In addition, we show that the variability of crop yields is expected to increase with increasing warming in many world regions. We provide, for each crop model and climate model, patterns of mean yield changes that allow for a simplified description of yield changes under arbitrary pathways of global mean temperature and CO$_2$ changes, without the need for additional climate and crop model simulations.

1 Introduction

Climate change exerts a substantial and direct impact on food security and hunger risk by altering the global patterns of precipitation and temperature which determine the location of arable land (Parry et al., 2005; Rosenzweig et al., 2014) as well as the quality (Müller et al., 2014) and quantity (Müller and Robertson, 2014; Lobell et al., 2012; van der Velde et al., 2012) of crops comprising most of the world food supply.

Climate change alone, *by itself*, climate change is expected to reduce global production of the four major crops wheat, maize, soy and rice on current agricultural areas (e.g., Rosenzweig et al., 2014; Challinor and Wheeler, 2008; Peng et al., 2004). Facing an increasing food demand due to population growth and economic development, these reductions will have to be overcompensated by 1) the direct physiological impacts of increased atmospheric CO$_2$ concentrations (Kimball, 1983), which are beyond local human control; as well as 2) advances in agricultural management (e.g. fertilizer input or irrigation), technology, and breeding (Jaggard et al., 2010) or 3) expansion of agricultural land (Frieler et al., 2015; Smith et al., 2010).

In conjunction with these long term changes, global warming is also expected to contribute to an increase in the frequency and duration of extreme temperatures and precipitation (droughts, floods, and heat waves), which may increase the near term variability of crop yields and trigger short term crop price fluctuations (Brown and Kshirsagar, 2015; Mendelsohn et al., 2007; Tadesse et al., 2014).

The emission anthropogenic emissions of greenhouse gases are expected to influence crop yields via several channels. On the one hand, the associated climate Cli-
mantic changes will modify the length of the growing season (Eyshi Rezaei et al., 2014), water availability, and heat stress ( Lobell et al., 2012; Müller and Robertson, 2014; Schlenker and Roberts, 2009); and on the other, higher concentrations of atmospheric CO₂ are expected to increase the water use efficiency in C₃ (e.g. wheat, rice, soy) and C₄ (maize) crops, and enhance the rate of photosynthesis in C₃ crops (Darwin and Kennedy, 2000). Global Gridded Crop Models (GGCMs) are particularly designed to account for these effects. They provide a complex process-based implementation of our current understanding of the mechanisms underlying crop growth, and are the primary tool for crop yield projections (e.g., Rosenzweig et al., 2014) which in turn are a prerequisite for assessing potential changes in prices (Nelson et al., 2014) and food security (Parry et al., 2005).

However, these process-based crop yield projections rely on spatially explicit realizations of the driving weather variables such as temperature, precipitation, radiation, and humidity, often at daily resolution, as provided by computationally expensive Global Climate Model (GCM) simulations. The GGCMs themselves also require significant computational capacity. These requirements generally limit the number and duration of emission scenarios that can be considered. The so-called pattern scaling approach is a well-established method to overcome these limits. Output from GCMs has been shown to be, to some extent, scalable to different global mean temperature (GMT) trajectories not originally covered by GCM simulations (Santer et al., 1990; IPCC-TGICA, 2007; Mitchell, 2003; Giorgi, 2008; Solomon et al., 2009; Frieler et al., 2012; Heinke et al., 2013) (Santer et al., 1990; Mitchell, 2003; IPCC-TGICA, 2007; Giorgi, 2008; Solomon et al., 2009; Frieler et al., 2012; Heinke et al., 2013). Scaled climate projections have also been used as input for different impact models (Oyebamiji et al., 2013; Stehfest et al., 2014) to gain flexibility with regard to the range of emission scenarios considered.

Building upon such a framework, we present a method to extend the capacity of crop yield impact projections by relating simulated crop yield changes to two highly aggregated quantities – global mean temperature change (ΔGMT) and atmospheric CO₂ concentration (pCO₂) – by means of simplified function. ΔGMT and pCO₂ are the standard output standard outputs of simple climate models, which allow for highly efficient climate projections for any emissions scenario by emulating the response of the complex GCMs (Meinshausen et al., 2011). Here “emulating” means that the simplified representation is designed to reproduce the complex model response of the complex model for the originally simulated scenarios but also allows for its inter- or extrapolation to other scenarios. In this way we Test to what extent crop yield changes, as one example of climate change impacts, can be described directly in terms of GMT and pCO₂ changes. Our approach is different from other emulators building on regional which use spatially explicit climate projections as input for the simplified functions emulating complex crop models’ responses to these forcings (Blanc, 2017) (Oyebamiji et al., 2015; Blanc, 2017). While these approaches only emulate the crop model responses of the complex crop model, the approach presented here implicitly provides a simplified description of both the GCMs’ regional patterns of climate change and the associated response of the crop models.

We test to what extent climate change impacts such as crop yields can be directly described in terms of GMT (and pCO₂) changes without an intermediate scaling of the regional climate changes. Such a direct description of the simulated impacts – in contrast to scaling the climate changes for specific emission scenarios and then using the scaled climate projections as input for impact model simulations – has the advantage of saving computation time. provides high computational efficiency, making the approach e.g. applicable within for example, in Integrated Assessment Models and even when no impact model is accessible. In principle, scaled but spatially explicit climate projections could also be used as input for spatial explicit crop model emulators (Blanc, 2017) to reach high efficiency. However, in this case the scaling of other emulators could be used in this setting, however requiring an additional step of first scaling the climate changes to the climate information has to be carefully adjusted to provide the kind of climate information required by the impact model impact emulator and this two-step approach also mean two approximations that may lead to higher deviations than the one-step approach proposed here specific emission scenario.

The emulator introduced here allows for multi-impact-model projections for arbitrary emission scenarios as long as crop-model ensemble projections are available for a limited set of scenarios. This offers a practical way of keeping track of a relevant but often-ignored source of uncertainty which is manifested in the considerable spread across different crop models and other process-based impact models (Rosenzweig et al., 2014; Schewe et al., 2014). This uncertainty is particularly critical when estimating socio-economic consequences (e.g., Nelson et al., 2014).

We test the approach using an ensemble of yield projections of the four major cereal crops (Crops maize, rice, soy, and wheat), generated within the first phase (“Fast Track”) of the Inter-sectoral Impact Model Intercomparison Project (ISIMIP, Warszawski et al., 2014). For a number of ΔGMT intervals we compare the spread in yield outcomes induced by the choice of emission scenario with that induced by the choice of GCM and GCM, respectively. A low scenario-induced spread means that GCM- and GGCMS-specific yield projections can be approximated by a simplified relationship with global mean temperature change without accounting for the underlying emission scenario, which is a prerequisite to applying the simplified relationship to other emission scenarios. The test is done at each grid cell point and separately for simulations of purely rain-fed yields and fully irrigated yields. Multi-model ensembles of crop yields over such a wide range of crops, CO₂ concentrations, and irrigation options are a new prospect and the ISIMIP data provides a uniquely broad suite of crop yield impact simulations encompassing output from five GGCMS, forced with output from up to five GCMs, and four Representative Concentration Pathways (RCPs, van Vuuren et al., 2011).
In Section 2 we describe the ISIMIP data and the methods used to test for scenario dependence and adjustment for different levels of CO₂ concentration at given global mean temperature at the respective GMT levels. We propose three methods to generate these patterns based on the available complex model simulations, and describe the related approaches to estimate GGCM- and GCM-specific yield changes for new ΔGMT trajectories not originally covered by GCM-crop-model simulations. In Section 4 we present a quantification of the projection errors as compared to actual simulations by the complex grided crop models. Finally, in Section 5 we quantify the residual inter-annual variance of the simulated crop yields in terms of global mean temperature change for each combination of ΔGMT change across all combinations of crop and climate models. Section 6 provides a summary.

2 Data and Methods

2.1 Crop yield simulations

We use projections from five different GGCMs (GEPIC, LPJ-GUESS, LPJmL, PEGASUS, and pDSSAT) that participated in the first simulation round of ISIMIP (Rosenzweig et al., 2014; Warszawski et al., 2014) in order to test for a dependence of projected yield changes on the global mean temperature GMT pathway (see Table 1 for their basic characteristics). Each crop model was forced by climate projections from five different GCMs (HadGEM2-ES, IPSL-CM5A-LR, MIROC-ESM-CHEM, GFDL-ESM2M, NorESM1-M) generated for four RCPs (RCP2.6, RCP4.5, RCP6.0, RCP8.5) in the context of the Coupled Model Intercomparison Project, phase 5 (CMIP5, Taylor et al., 2012). CMIP5 was an effort by the climate modelling community to provide a new suite of climate simulations in time for the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5). The RCPs cover the range from climate mitigation (RCP2.6, RCP4.5) to business-as-usual (RCP6.0) and high emissions scenarios (RCP8.5). Climate projections have been bias-corrected to better match observed historical averages of the considered climate variables. For the future, the bias-correction preserves absolute changes in monthly temperature and relative changes in monthly values of the other variables simulated by the GCMs while also correcting the daily variability about the monthly mean (Hempel et al., 2013). Separate simulations are available for each of the four major crops: wheat, maize, rice, and soy, on a global 0.5 x 0.5 degree grid, covering the time period from 1971–2099. The considered crop is assumed to grow everywhere on the global land area, only restricted by soil characteristics and climate but independent of present or future land use patterns (“pure crop” simulations). Each model has provided a pair of simulations (“rums”) for each climate change scenario: 1) a rain-fed run and 2) a full-irrigation run assuming no water constraints. This design provides full flexibility with regard to the application of future land use and irrigation patterns. While the “default” crop yield simulations (Ŷ_{noCO2}Ŷ_{fixedCO2}) account for the fertilization effects due to the increasing levels of atmospheric CO₂, the ISIMIP setting also includes a sensitivity experiment where the impact models were forced by the same climate change projections from HadGEM2-ES. RCP8.5 but CO₂ concentration was kept fixed at a “present day” reference level that differs from GCM to GCM (see Table 1). We will refer to this run as “fixed CO₂” run and indicate the associated crop yields by Ŷ_{noCO2}Ŷ_{fixedCO2}. As a special case, the “default” simulations for pDSSAT do not use annual CO₂ changes. Instead, CO₂ was changed every 30 years using the average CO₂ of the respective 30 year time slice.

2.2 Effect of temperature change

We analyse the dependence of yield changes on ΔGMT separately for rain-fed and full-irrigation simulations, and for each crop are considered separately. Considering e.g., wheat yield changes under full irrigation, Since the timing of global warming differs between GCMs and scenarios, we group all available data into ΔGMT intervals (bins) separated by 0.5°C steps with 0.5°C width (±0.25°C around the central temperature), where ΔGMT is relative to the present day (1980–2010 average) reference level. For all annual data falling into a given interval and at one specific grid point we apply a separate one-way analysis of variance (ANOVA fixed effects model) to individually calculate the variance explained by 1) different GGCMs, 2) the GCMs, and 3) the RCPs. The quantification of the RCP-dependence of the relationship between global warming and yield changes is limited to a warming range up to 2 to 3°C above present depending on the GCM because only one RCP (RCP8.5) reaches temperatures above this threshold. However, we also provide the patterns of yield changes for the higher concentration scenario. In the main text, all figures except Figure 1 & 10 refer to a ΔGMT level of 2.5°C (see Figure 1 for the associated years included) but the Supplement contains the figures for the other levels, and all figures except 3, 4, and 11 refer to crop model simulations driven by HadGEM2-ES climate. See Figure 1 for the years asso-
Table 1. Basic crop model characteristics with respect to 1) the implementation of CO₂ fertilization effect (as affecting radiation use efficiency (RUE), transpiration efficiency (TE), leaf level photosynthesis (LLP), or canopy conductance (CC)), 2) the accounting for nutrient constraints with respect to the CO₂ fertilization effect and associated assumption with respect to fertilizer application (N = nitrogen, P = phosphorus, K = potassium), 3) implemented adaptation measures, and 4) starting conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>CO₂ fertilization</th>
<th>Fertilizer use</th>
<th>Nutrient limitation</th>
<th>Adaptation</th>
<th>Starting conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEPI</td>
<td>RUE, TE pCO₂ of the fixed CO₂ run: 364 ppm</td>
<td>Limitation of potential biomass increase due to N stress (flexible N application based on N stress &gt;10% up to an upper national application limit according to FertiStat (FAO, 2007)).</td>
<td>Fixed present-day FAO FertiStat database (FAO, 2007), fixed present-day P application rates following FAO FertiStat database (FAO, 2007)</td>
<td>FertiStat</td>
<td>decadal adjustment of planting dates (incl. switch between winter and spring wheat); total heat units to reach maturity remain constant; decadal adjustment of winter and spring wheat sowing areas based on temperature present day</td>
</tr>
<tr>
<td>LPJ-GUESS</td>
<td>LLP, CC pCO₂ of the fixed CO₂ run: 379 ppm</td>
<td>No consideration of spatial and temporal changes in soil nutrient limitation</td>
<td>Soil nutrient limiting factors are not accounted for</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPJmL</td>
<td>LLP, CC pCO₂ of the fixed CO₂ run: 370 ppm</td>
<td>Soil nutrient limiting factors are not accounted for</td>
<td>Soil nutrient limiting factors are not accounted for</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEGASUS</td>
<td>RUE, TE pCO₂ of the fixed CO₂ run: 369 ppm</td>
<td>Fixed N, P, K application rates (IFA, 2002)</td>
<td>Fixed present-day application rates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pDSSAT</td>
<td>RUE, LLP, CC pCO₂ of the fixed CO₂ run: 330 ppm</td>
<td>Fixed N present-day application rates</td>
<td>No adjustment of planting dates; total heat units to reach maturity remain constant</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.3 Effect of pCO₂ change

The direct effect of CO₂ fertilization on crop yields is expected to introduce some scenario dependence in the relationship between GMT change and yield change. We test to what degree the scenario dependence of the relationship can be explained by introducing atmospheric CO₂ levels pCO₂ as an additional predictor for within-bin fluctuation of yields. To this end, we evaluate two different approaches to estimate the direct CO₂ CO₂ effect on crop yields within the different GMT bins, described in detail in Section 3.2 below. The two approaches differ in terms of the crop model simulations that they require: approach (a) only requires the default crop yield simulations with increasing pCO₂ whereas approach (b) requires a pair of simulations with increasing pCO₂ and with fixed pCO₂ at present-day reference level.

To evaluate and compare the performance of the two approaches we consider large scale regional average yields based on fixed present day (1998-2002) land use and irrigation patterns from MIRCA2000 (Pottmann et al., 2010) and assess the reproducibility of the original RCP2.6, RCP4.5.

2.3.1 Approach (a)

For all years falling into a specific ΔGMT bin, approach (a) fits the following linear regression model to the response

\[ Y = \alpha + \beta_1 \Delta GMT + \beta_2 pCO₂ + \epsilon \]

where \( Y \) is the yield, \( \Delta GMT \) is the change in global mean temperature, \( pCO₂ \) is the atmospheric CO₂ concentration, and \( \epsilon \) is the error term.
of yields in the default simulation to the increase in pCO₂:
\[
\Delta Y_{\text{varCO₂}}(i, t) = \Delta Y_{\text{clim}}(i) + a_1(i) \cdot (p\text{CO₂}(t) - 370\text{ppm}) + \epsilon(i, t),
\]
(1)
where \(\Delta Y_{\text{varCO₂}}(i, t)\) is the absolute yield change in grid point \(i\) and year \(t\) with respect to the historical reference period (1980–2010) and \(p\text{CO₂}(t)\) is the atmospheric CO₂ concentration of the corresponding year. In this statistical model, \(\Delta Y_{\text{clim}}(i)\) represents an estimate of the purely climate-induced yield change at the respective bin temperature, but assuming a fixed year-2000 pCO₂ of 370 ppm (i.e. without CO₂ fertilization), \(a_1(i)\) represents the added effect of CO₂ fertilization, and \(\epsilon(i, t) \sim N(0, \sigma^2)\) represents the residual error.

2.3.2 Approach (b)

Approach (b) fits the following linear regression model to the yield difference between the default and fixed-CO₂ simulation for all years falling into a specific ΔGMT bin:
\[
Y_{\text{varCO₂}}(i, t) - Y_{\text{fixedCO₂}}(i, t) = a_1(i) \cdot (p\text{CO₂}(t) - p\text{CO₂}_{\text{ref}}) + \epsilon(i, t),
\]
(2)
where \(Y_{\text{varCO₂}}(i, t)\) and RCP6.0 projections based on the emulated yield patterns (section 4) \(Y_{\text{fixedCO₂}}(i, t)\) is the absolute yield in grid point \(i\) and year \(t\) of the default and fixed-CO₂ simulation, respectively, \(p\text{CO₂}(t)\) is the atmospheric CO₂ concentration of the default simulation during the respective year and \(p\text{CO₂}_{\text{ref}}\) is the crop-model specific pCO₂ value of the fixed-CO₂ simulation (see Table 1). In this statistical model, \(a_1(i)\) represents the CO₂ fertilization effect and \(\epsilon(i, t) \sim N(0, \sigma^2)\) represents the residual error. No intercept is estimated in this model because yields from the default and fixed-CO₂ runs are expected to be identical if \(p\text{CO₂}(t) = p\text{CO₂}_{\text{ref}}\). The purely climate-induced yield change at a fixed year-2000 pCO₂ of 370 ppm \(\Delta Y_{\text{clim}}(i)\) can then be derived as:
\[
\Delta Y_{\text{clim}}(i) = \Delta Y_{\text{fixedCO₂}}(i) + a_1(i) \cdot (p\text{CO₂}_{\text{ref}} - 370\text{ppm}),
\]
(3)
where \(\Delta Y_{\text{fixedCO₂}}(i)\) is the average yield change in the respective warming bin of the fixed CO₂ simulation with respect to the historical reference period and \(a_1(i) \cdot (p\text{CO₂}_{\text{ref}} - 370\text{ppm})\) corrects for the different pCO₂ref used by each GGCIM.

2.4 Emulator of temperature and CO₂ effects

Based on the spatial patterns of purely climate-induced yield change \(\Delta Y_{\text{clim}}(i)\) and added CO₂ fertilization effect \(a_1(i)\), which are derived separately for each rain-fed and irrigated crop and specific to each crop model and GCM, we propose the following two-step interpolation method to compute crop yield changes for any given pair of ΔGMT and pCO₂, using either the coefficients from approach (a) or (b):

1. linear interpolation of \(\Delta Y_{\text{clim}}(i)\) between the two neighbouring warming bins to the desired ΔGMT value,
2. addition of the CO₂ pattern described by \(a_1(i) \cdot (p\text{CO₂} - 370\text{ppm})\), where \(a_1(i)\) is also interpolated linearly between the respective coefficients from the neighbouring warming bins.

The application of these two steps using coefficients from method (a) above will be called emulator approach (a); their application using coefficients from regression method (b) will be called emulator approach (b). In addition, we propose a third, very basic emulator approach (c) where the yield change for any given ΔGMT is derived from a simple linear interpolation of the average yield change in the neighbouring warming bins of the default simulations \(\Delta Y_{\text{varCO₂}}(i)\) with respect to the historical reference period, without using the associated pCO₂ as additional predictor.

The linear interpolation of any of the previous coefficients between two neighbouring warming bins is illus-
treated for a ∆GMT of 2.3°C as follows:

\[
\text{coef}(i, 2.3°C) = (1 - \delta) \cdot \text{coef}(i, 2°C) + \delta \cdot \text{coef}(i, 2.5°C),
\]

where \( \text{coef} \) can be \( \Delta Y_{\text{clim}}(i), a_1(i), \) or \( \Delta Y_{\text{varCO2}}(i). \)

Using GGCM projections for the HadGEM2-ES climate input to train the emulators, we test which of the emulator approaches, (a), (b) or (c), provides the best reproducibility for yield changes simulated under the four RCPs (Section 4). While approach (b) requires a pair of crop model simulations – one with time-varying \( pCO_2 \) and one with fixed present-day \( pCO_2 \) – approach (a) and (c) only require the default simulations with time-varying \( pCO_2 \).

Thus, a comparison of the three approaches could provide some important guidance regarding future crop model experiments required to allow for the proposed highly efficient emulation of crop model simulations. Since simulated crop yields are subject to considerable inter-annual variability, we also test what effect the amount of available training data has on the reliability of the derived regression coefficients. For that purpose, we train the emulators using either all available simulation data from the four RCPs or only simulation data from RCP8.5 and compare the fraction of the land surface for which derived fits are statistically significant as well as the difference between simulated and emulated yield changes. Due to the 30-year time slices of constant \( pCO_2 \) used by pDSSAT in the default run, approach (a) cannot be applied to this model using only RCP8.5 data. Since only RCP8.5 reaches \( \Delta \text{GMT} > 3.5°C \) this limits the temperature range of emulator approach (a) for pDSSAT even when using all available training data.

We evaluate and compare the performance of the three emulator approaches at the grid scale as well as the scale of large regions. Grid point yields (in t/ha) are multiplied by the fixed year-2000 crop-specific growing area from the MIRCA2000 dataset (Portmann et al., 2010) to derive regional total crop production (in t). MIRCA2000 provides gridded growing areas for a total of 26 rain-fed and irrigated crops based on a combination of census, remote sensing and other geographic data sources.

3 Mean Yield Change with Global Mean Temperature Change

3.1 Patterns of relative changes at different levels of global warming and main sources of variance

In general, increasing global mean temperatures correspond to an expansion of arable land to higher latitudes with concurrent yield reductions in equatorial regions. The highest positive changes in projected yields under rain-fed conditions at 2.5°C ∆GMT are typically in the northern high latitudes and mountainous regions for all crops (Figure 2 for wheat, figures for other crops in the Supplement). These locations were previously inhibited by a short growing season, which extends with increasing air temperature (Ramankutty et al., 2002). Yield gains also occur over previously moisture limited regions, such as the northwestern U.S. and north-eastern China, in agreement with the findings of Ramankutty et al. (2002). In contrast, near the equator most crop yields decrease, especially maize and wheat. Since most cultivated land currently lies in low and middle latitudes, potential yield changes in those regions contribute a higher relative importance for today’s food production system than changes in high latitudes.

While variations exist in the magnitude of projected yield changes, there is a high degree of consistency in the direction of yield change across ensemble members, especially over the high latitudes, where most of the largest projected yield changes occur, but where yields are in general smaller (Figure 3; Figure 4). Utilizing output from all available combinations of one GCM, GGCM, and RCP scenario, more than three-quarters of the ensemble members indicate increasing crop yields over the upper mid latitudes in the northern hemisphere for all crops at 2.5°C.

The simulated yield values at each grid point and within each GMT bin are subject to variation due to the selection of impact model, GCM forcing, and emissions scenario. When considering all of these factors, the variance attributable to the impact model selection is much greater than that associated with the GCM or scenario choice in most regions (Figure 4; Figure 4). This holds for both rainfed and irrigated simulations and at all global mean warming bins above 1°C. The predominance of the impact model component in total variance is particularly evident in the middle to high latitudes for all four cereals considered, where impact model variance accounts for up to 90% of the grid point variance at 2.5°C.

3.2 Direct impacts of increasing \( pCO_2 \)

In addition to air temperature warming, \( pCO_2 \) has a direct influence on crop yields. As it varies within the different ∆GMT bins, it is expected to induce part of the fluctuations of the yield changes at given GMT levels. We find that this CO2 effect is not scenario dependent (see Figure 5; CO2 effect shows little scenario dependence (see Figure 5 for the global average effect within the LPJml simulations at ∆GMT=2.5°C), consistent with a short response time of plants to \( pCO_2 \) changes. As expected, the \( CO_2 \)-induced yield differences increase with heightened atmospheric \( CO_2 \) level under all emissions scenarios, implying a stronger \( CO_2 \) fertilization impact with increased \( pCO_2 \). A least squares fit to the yield differences versus greenhouse gas level within each ∆GMT bin allows for a quantification of the direct \( CO_2 \) effect at each level of warming based on global \( pCO_2 \), rather than the emissions pathway. The underlying assumption is that the effect of the temperature variation within the 0.5°C range of each ∆GMT bin will be minimal compared to the effect of the \( CO_2 \) variation across all RCPs \( pCO_2 \).

To quantify the extent of the \( CO_2 \) induced scenario dependence and its potential reduction at each grid point, we use two methods to determine the \( CO_2 \) effect on crop
Figure 2. Average potential wheat yield change at $\Delta$GMT=$2.5^\circ$C as a percentage of the mean historical yield (1980–2010 average) under rain-fed conditions for each crop model forced by HadGEM2-ES. The average is calculated across all RCPs which reach the global mean warming interval from 2.25 to 2.75$^\circ$C, namely RCP4.5, RCP6.0, and RCP8.5. Note that pDSSAT is run over a limited domain excluding areas north of 60$^\circ$N. Regions with marginal historical yields (defined as lying below the 2.5% quantile of historical yields on year-2000 cropland) are masked to avoid exaggerated relative yield increases. Analogous figures for different crops, for irrigated conditions, as well as for absolute yield change (in t/ha) are available as supplementary online material in the Supplement.
regions where both approaches disagree regarding the direction of change, such as for rice where there is disagreement on the direction of yield changes in southeast Asia. In relative terms (estimated climate change-induced yield change divided by simulated present-day yield), both methods show very large values of frequently alternating sign in areas such as the Arabian peninsula or the northern Sahel (Fig. 6, upper panel). This is likely due to the very low present-day yield potential in these regions, leading to division by values close to zero. In the regional evaluation of the different emulator methods below, we will account for these regional differences in baseline yields by weighting potential yield changes by present-day growing areas. High latitudes of both Western North America and Eastern Russia for wheat and parts of Southeast and South Asia for all crops. Patterns of climate-induced yield change match better between both approaches under irrigated conditions. In GEPIC, both approaches disagree on the direction of change for maize yields over large parts of Europe. In LPJ-GUESS, both approaches disagree on the direction of change in most of the tropics for all crops. While tropical yield change is predominantly negative in approach (b) mirroring results of the other crop models, approach (a) estimates mostly positive climate effects on tropical crops. In pDSSAT, approach (a) generally produces larger areas with negative yield change than approach (b). At the same time, positive yield effects in approach (a) have a larger magnitude than those in approach (b) in many regions. In PEGASUS, both approaches disagree on the direction of change over large parts of the U.S. for maize and soybeans, and large parts of China for wheat.

The estimates of CO2-induced yield changes CO2-induced yield change also differ between the two methods (Figure 7). Method (b) results in a positive CO2 effect in most regions, except for some low-yielding areas and the potentially important cases of soybean in southern and eastern South America, and rice in north-west India and Pakistan, where it results in a negative effect of rising pCO2 on yield. With method approaches (Figure 7 for LPJmL results under rain-fed conditions). We expect CO2 fertilization to have a positive or at least neutral effect on yields, and this is confirmed by approach (b) for all GGCMs and crops. Only GEPIC simulations show negative CO2 effects on soybean and wheat yields in a few regions for approach (b). This can be explained by nutrient interactions in the model: CO2 fertilization leads to yield increases first but also increases nutrient depletion in the soil compared to the fixed-CO2 run. If fertilizer application is insufficient to replenish nutrient stocks this can lead to lower yields despite the beneficial effect of higher pCO2. With approach (a), on the other hand, areas of negative estimated CO2 effect are much more widespread, and generally CO2 effects are widespread in all GGCMs and crops. Generally, the magnitudes of the estimated CO2 effect are again much larger with method CO2 effect are also much larger, often surpassing those of approach (b). As a preliminary conclusion, the results obtained with method (b) for the separate even in regions where the direction of change matches. Given that approach (a) contradicts our expectation of how CO2 fertilization should affect yields in many regions we conclude that approach (a) is not reliable in separating the effects of climate change and pCO2 change on potential yields appear more realistic than those obtained with method (a).
Figure 4. Fraction of total variance attributable to the impact models (GGCMs, left), climate models (GCMs, middle), and scenarios (RCPs, right) for each crop. Figure shown for rain-fed runs at $\Delta GMT = 2.5^\circ C$; an analogous figure for irrigated runs is provided as supplementary online material in the Supplement.

4 Validation of three emulator approaches

Based on the on yield from those of pCO$_2$ change. By design, climate-induced patterns (assuming fixed year 2000 levels of CO$_2$) of relative yield changes and the associated within-bin relationship between CO$_2$ and crop yields identified in section 3, we propose the following two-step interpolation method to compute crop yield changes for any given pair of $\Delta GMT$ and pCO$_2$, using either of the above regression methods (a) or (b): Linear interpolation between the temperature-specific, CO$_2$-adjusted yield patterns of neighboring $\Delta GMT$ bins ($a_0(\Delta GMT)$ from method and CO$_2$-induced yield changes add up to the full yield change (see Equation 1) which is why the difference between the patterns of estimated CO$_2$ effect explains why climate-change patterns from Figure 6 also differ substantially between both approaches in some regions. Approach (a) or $Y_{\Delta GMT} = Y_{\Delta GMT} + a_1(\Delta GMT)$ from method (b) to the desired $\Delta GMT$ value. Addition of the CO$_2$ pattern described by $a_1$ has a structural disadvantage to approach (b) in that it estimates both the climate-induced and CO$_2$-induced effect on yields from the same linear regression model (Equation 1). Besides changes in pCO$_2$ annual yields in each warming bin are subject to substantial inter-annual climate variability which means that individual years with a higher pCO$_2$ do not necessarily have a higher yield. In contrast, approach (b) only estimates the $\text{CO}_2^{-370}$ppm, where the pattern of scaling coefficients $a_1$ is also interpolated linearly between the scaling coefficients from neighboring temperature bins. The application of these two steps using regression method (a) will be called emulator-induced yield change from the regression model (Equation 2) while both the default and the fixed-CO$_2$ run are subject to identical climate variability. There is inter-annual variability in the CO$_2$-induced yield change as well (see Figure 5 for the global average effect), however, it is much smaller than the total yield variability. While approach (a) : their application using regression method (b) will be called emulator approach and (b). In a third, very basic emulator approach (c), crop yield change patterns for a given should provide similar estimates of the CO$_2$-induced yield change given a large sample, our sample size is limited by the number of years falling into each $\Delta GMT$ level are derived from an interpolation between the two neighboring $\Delta GMT$ bins’ average patterns; where these average patterns are derived from the bin (Table 2). This number varies between seven years in the 4.5 and 5.0$^\circ C$ bin and up to 66 years in the 1.0$^\circ C$ bin when yield data from all RCPs are used to train the emulator. The number of years varies between seven and 13 years if only data from RCP8.5 projections of the individual climate and
Validation of three emulator approaches

Using GGCM projections for the HadGEM2-ES climate input, we test which of the approaches, (a), (b) or (c), provides the best reproducibility for RCP2.6, RCP4.5, and RCP6.0 when estimates of the climate-induced and CO2-induced effects are based on RCP8.5 projections. While approach (b) requires a pair of crop model simulations – one with time-varying pCO2 and one with fixed pCO2, approach (a) only requires the default simulations with time-varying pCO2. Approach (c) assumes that yield changes can be estimated using only four RCPs. For that purpose, we apply each emulator with time series of ΔGMT as a predictor without consideration of the associated pCO2. Thus, a comparison of the three approaches could provide some important guidance regarding future crop model experiments required to allow for GMT and pCO2 from the RCPs and compare emulated yield changes in each grid point and as well as total crop production for 10 large world regions to those simulated by

Figure 5. Difference in global mean yield change (sum of rain-fed and irrigated, and weighted by present-day year-2000 growing areas) between the default (Y_{CO2-varCO2}) and fixed CO2 simulations (Y_{noCO2-fixedCO2}), for each crop over the range of pCO2 associated with the ΔGMT=2.5°C bin. Results are as simulated by LPJmL forced with output from HadGEM2-ES. Each color represents an emission scenario and black. Points mark individual years while dotted lines and shaded areas indicate the linear best fit and its 95% confidence interval for each crop. The black dotted line indicates the linear best fit through all available scenarios. Analogous figures for other GGCMs and warming bins are available in the Supplement.

crop model simulations accounting for the CO2 fertilization effect. E.g. to derive the crop yield change pattern for a global mean warming of 2.5°C are used. Given the limited sample size and possibly large variability, the derived fits are often not statistically significant. For approach (a) we found that derived fits were rarely significant on more than 25% of the crop-specific growing area (Pörtmann et al., 2010) using a p-value of 0.05 (figure available in the Supplement). Values were even lower if only RCP8.5 was used for the regression. In contrast, fits derived by approach (b) were mostly statistically significant (p < 0.05) on more than 70% of the growing area, often on more than 90% of the area. We also found only a small negative effect in terms of statistical significance if only RCP8.5 was used in approach (b).
Figure 6. Climate change-induced yield changes at $\Delta GMT=2.5^\circ C$ of global warming and year 2000 pCO$_2$ level (370 ppm). Left column: Patterns of $\Delta Y_{\text{clim}}(i)$ derived at each grid point $i$ by approach (a) (see Equation 1). Right column: Corresponding patterns of $\Delta Y_{\text{clim}}(i)$, derived by approach (b) (see Equation 3). Both types of patterns are derived from LPJmL simulations forced by HadGEM2-ES assuming rain-fed conditions and expressed as absolute differences compared to the historical period (1980–2010). Rows: Different crop types. Analogous figures for irrigated conditions, for different GGCMs, and using relative instead of absolute yield changes are available in the Supplement.

Table 2. Number of years of yield data available in each $\Delta GMT$ bin for HadGEM2-ES. Only RCP8.5 reaches warming levels above $3^\circ C$.

<table>
<thead>
<tr>
<th>Data used</th>
<th>$\Delta GMT$ bin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5$^\circ C$</td>
</tr>
<tr>
<td>all available scenarios</td>
<td>47</td>
</tr>
<tr>
<td>RCP8.5 only</td>
<td>10</td>
</tr>
</tbody>
</table>
Figure 7. CO$_2$-induced yield changes at 2.5°C of global warming for LPJmL forced by HadGEM2-ES assuming rain-fed conditions. Analogous to Figure 6, but showing the scaling coefficients $a_1(i)$ from approach (a) (left column) and approach (b) (right column), multiplied by the average pCO$_2$ change compared to year 2000 (370 ppm) across all years falling into the GMT bin. Rows: Different crop types. Analogous figures for irrigated conditions, for different GGCMs, and using relative instead of absolute yield changes are available in the Supplement.
Figure 8. Validation of the three emulator approaches. Maps show the difference (emulated minus simulated) between the simulated LPJmL yields forced by HadGEM2-ES climate for RCP4.5 under rain-fed conditions for all years falling into the ∆GMT bin of 2.5°C (2066–2094) and the emulated yields for the same years based on approach (a) (left column), approach (b) (middle column), and approach (c) (right column). Rows: Different crops. MAD: mean absolute difference, regardless of sign, averaged across all grid points. Analogous figures for irrigated conditions and for different GGCMs are available in the Supplement.

Figure 8 shows results for the LPJmL model, when applying the emulators trained on all available data to reproduce rain-fed yields under RCP4.5. Figures for other RCPs, irrigated yields and other GGCMs are available in the Supplement.

Approach (a) generally leads to the largest differences relative to the simulated yield change (Fig. 8, left column). In particular maize, rice, and soybean yields are underestimated for much of North America, and overestimated in Europe and South America, temperate South America, and Australia. Wheat yields are overestimated, e.g., in Canada. These discrepancies are mainly due to the climate-change effect estimated by approach (a) (cf. Figure 6), whereas the CO2 fertilization effect even points in the opposite direction in many of these regions (cf. Figure 7). In fact, we note again that approach (a) estimates the CO2 fertilization effect to be negative in some regions (Figure 7), which is not consistent with theory and empirical evidence, in Canada.

Approach (b) also leads to some substantial deviations from the potential yields simulated by LPJmL, in percentage terms, mainly in the northern hemisphere and in Australia (Figure 8, top panel). But large relative differences are mainly found outside the major growing regions of the respective crop, in areas where absolute potential yields are low today. Correspondingly, absolute differences between the LPJmL simulations and the emulator Spatial patterns of over and underestimation are broadly similar to approach (a), but the magnitude of the difference is generally slightly lower. In the tropics, approach (b) are modest (Figure 8, bottom panel, middle column). An important exception is the underestimation of simulated maize and rice yields in southern North America. We note that LPJmL itself has limitations in simulating yield variability in this region (Frieler et al., 2017). Often leads to a higher deviation from the simulated yields than approach (a), particularly for rice and soybeans in South America.

Finally, approach (c) leads to a similar pattern of deviations from the simulated yield potential yields as approach (b), but with a slightly smaller magnitude (Figure 8, right column). Thus, considering overall performance at the grid point level for this particular case (2.5°C warming under RCP4.5), the simple approach (b) is recommended.
approach (c) often leads to an overestimation of yields whereas approach (b) tends to underestimate simulated yields. The average deviation between emulated and simulated yields (designated as MAD in Figure 8) is similar for approach (b) and (c). Approach (c) performs slightly better than approach (b) for rice, and both approach (b) and (c) perform better than approach (a) for all four crops. Differences between the three emulators are smaller when reproducing RCP6.0 and RCP8.5 (figures available in the Supplement).

The difference between emulator approach (c) produces results which are closest to the LPJmL simulation. b) and (c) is even smaller in the other crop models than in LPJmL (figures available in the Supplement). Overall, MAD between emulated and simulated yields is up to 50% higher than LPJmL in PEGASUS, roughly twice as high in GEPIC and up to three times as high in pDSSAT. In LPJ-GUESS, MAD between emulated and simulated yields is similar for all three emulator approaches, even though the spatial patterns of over and underestimation differ.

Using only RCP8.5 instead of all available data to train the emulators has a detrimental effect on the performance, especially for approach (a). MAD between emulated and simulated yields increases by a factor of more than three, even close to four for some GGCMs and crops, under RCP4.5. MAD for approach (b) and (c) also increases by a factor of more than two, although not as sharply as for approach (a) (figures available in the Supplement). Performance loss is lower for RCP6.0, with MAD generally less than twice as high. The emulator trained on RCP8.5 alone shows better performance in emulating RCP8.5 simulated yields than the emulator trained on all available data.

To get a more comprehensive indicator indication of the performance of the emulator for the whole 95-year time series (instead of just the 2.5°C bin) we use all three approaches to reproduce the simulated changes in crop production under RCP2.6, RCP4.5, and RCP6.0, and RCP8.5, as derived for 10 large scale world regions (cf. Figure 2 in Lotze-Campen et al. (2008) see Figure 10 for a map of the regions). Grid-point yields are aggregated to the regions assuming fixed year-2000 land use and irrigation patterns. Compared to potential grided yields, using production gives less weight to areas where a crop is not currently grown. The climate-induced and CO2-induced patterns of change were derived from RCP8.5, and we used the RMSE between the relative changes in crop production derived from the original simulations and their emulated counterparts across the other three scenarios. Since none of the emulators is expected to capture the relatively large inter-annual variability of simulated yields we compare simulated and emulated decadal production and calculate the RMSE over all decades of the relative difference between emulated and simulated decadal production (in %) as a measure of the performance of the emulator.

Of the two approaches that estimate warming and CO2-induced effects separately, approach (b) generally provides a better performance than approach (a) (see Figure 9 for LPJmL; Table 1 and supplementary online information Table 3 and the Supplement for all crop models). Performance of all emulator approaches varies substantially between regions. There are also considerable differences between crop models. For LPJmL, emulator approach (b) generally provides marginally better performance for many regions than approach (c) when emulating RCP2.6 and RCP4.5. This advantage of approach (b) is not found in . However, this is not consistent across the emulators for the other crop models. Taking into account that approach (b) requires additional crop model simulations with fixed CO2 and that performance is mostly very similar for approach (b) and (c), the very basic interpolation approach (c) may appear to provide the best compromise between emulator performance and complexity.

While none of the emulators is expected to capture the relatively large inter-annual variability of simulated yield changes, approach (c) allows to emulate the regionally averaged response of the process-based crop models to climate forcing estimated for RCP2.6, RCP4.5, and RCP6.0 (Fig. 10 for maize yields from LPJmL forced by HadGEM2-ES; analogous figures for other combinations are available as supplementary online material). Note though that the average deviation difference between emulated and simulated yields over the full 95-year time series is sometimes larger than the simulated yield production change in 2091–2099, especially in the low warming scenarios (marked by red crosses in Fig. 9 Figure 9). Table 3 compares the RMSE between emulated and simulated crop production in the largest producing region of each crop for all five crop models.

Figure 10 illustrates the performance of emulator approach (c) in reproducing decadal maize production as simulated by LPJmL forced by HadGEM2-ES. Emulated yields generally follow the simulated trends, although large errors exist, e.g., in North America, which also stands out in Figure 9 and Figure 8. Analogous figures for all crops, emulator approaches and crop models are available in the Supplement.

Similar to the grid point results, using only RCP8.5 to train the emulators leads to a performance loss for all emulator methods and all RCPs except RCP8.5. This performance loss is larger for approach (a) than approach (b) and (c), and is generally highest for RCP4.5 (figures available in the Supplement).

### 5 Increases in Regional Crop Yield Variance

In addition to estimating the yield changes change associated with a rise in average temperature, it is important to consider the implications of rising variance. Climate change is expected to increase not only the average temperature, but to impact the variance of temperature and precipitation, including an increase in the frequency and duration of extreme events. For this reason, when deriving simplified relationships between potential yields yield change and global climate change, it is crucial to account not only for the mean effects of rising temperature, but also their concurrent implications.
Figure 9. Root mean square difference (in %) between emulated and simulated regional decadal production (yields multiplied by year-2000 growing areas, combined for irrigated and rain-fed crops) for LPJmL forced by HadGEM2-ES climate projections. The emulator was built using all available data and used to reproduce yield changes in all four RCPs. For comparison, point symbols illustrate the average simulated yield change for 2091–2099 (same horizontal axis), using red crosses or blue circles depending on whether the error between emulated and simulated production is larger or smaller than the simulated change. Analogous figures for the other crop models are available in the Supplement.
Table 3. Average root mean square deviation difference between emulated and simulated decadal production (expressed in % of the simulated production) as in Fig. 9 in the largest producing region of each crop, for all five crop models forced by HadGEM2-ES climate projections. Average over all four RCPs. The values for all combinations of models, crops, and regions, and separately for each RCP, can be found in the supplementary online material Supplement. Top: emulators trained on all available data; bottom: emulators trained on RCP8.5 only.

(a) Emulators trained on all available data

<table>
<thead>
<tr>
<th>Model</th>
<th>Wheat, Europe</th>
<th>Rice, South Asia</th>
<th>Maize, North America</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a  b  c</td>
<td>a  b  c</td>
<td>a  b  c</td>
</tr>
<tr>
<td>LPJ-GUESS</td>
<td>2.242 2.254 2.213</td>
<td>4.033 2.163 3.729</td>
<td>5.870 5.466 5.359</td>
</tr>
<tr>
<td>LPJmL</td>
<td>1.777 1.768 1.596</td>
<td>2.582 2.371 1.786</td>
<td>6.923 5.494 5.846</td>
</tr>
<tr>
<td>PEGASUS</td>
<td>6.061 4.908 4.937</td>
<td>n.a. n.a. n.a.</td>
<td>8.762 8.533 8.496</td>
</tr>
</tbody>
</table>

(b) Emulators trained on RCP8.5 only

<table>
<thead>
<tr>
<th>Model</th>
<th>Wheat, Europe</th>
<th>Rice, South Asia</th>
<th>Maize, North America</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a  b  c</td>
<td>a  b  c</td>
<td>a  b  c</td>
</tr>
<tr>
<td>LPJmL</td>
<td>3.814 2.272 2.293</td>
<td>4.247 2.954 3.040</td>
<td>11.954 5.783 5.838 5.950</td>
</tr>
<tr>
<td>pDSSAT</td>
<td>4.863 n.a. 4.053</td>
<td>5.232 n.a. 4.230</td>
<td>8.065 n.a. 8.290 7.984</td>
</tr>
</tbody>
</table>

(1) Emulator approach (a) for pDSSAT only covers warming up to 3.5°C, i.e. up to 2070 under RCP8.5.

Figure 10. Comparison of simulated and emulated time series of regionally averaged aggregated crop production changes for LPJmL forced by HadGEM2-ES climate projections. Regional averages are calculated based on fixed present day land use and irrigation patterns. Results are shown for Maize and emulator approach (c).
for crop yield variance. Interannual yield variance can be computed for the same 0.5 °C warming bins as used above for the average yields, which we do here for all four major crops under the “no irrigation” scenario. To account for the variability across scenarios and models which is attributable to direct CO2 effects, the variance is calculated separately for the years of each RCP-GCM-GGCM combination over the historical period (1980–2010) period.

The global figures show broadly similar patterns across all four crops: Increases in yield variability in much of the northern hemisphere, particularly in North America, central Asia, and China; as well as in the southern mid-latitudes (Figure 11 for 2.5 °C GMT) of the CMIP5 ensemble based primarily on data availability. A majority of model combination projects decreasing variability in tropical regions (except for rice) as well as parts of Eastern Europe; but nowhere do more than 75% of the model combinations agree on a decrease in variability. In several instances increased variability occurs in highly productive regions such as in China for rice and the US, Brazil, and Argentina for soy. Wheat also has an increased variability in more than 75% of the crop model simulations over the highly productive regions in China and the U.S. Such an increase in variability, if realized, could manifest as impacts on the price, whose volatility is tightly linked to rapid changes in supply (Gilbert and Morgan, 2010).

6 Summary

Evaluating the impacts of climate change at different levels of global warming, and thus evaluating mitigation targets, requires a functional link between GMT and regional impacts. Here we have shown that changes in crop yields, as simulated by gridded global crop models, can be reconstructed based on GMT, with some limitations. The small spread of simulated yield change across the RCP scenarios as compared to the GGCMs and impact models implies that projected impacts at different GMT levels are not substantially dependent on the choice of emissions pathway. In this context, it has to be noted that the scenario setup of the ISIMIP crop model simulations was chosen specifically to minimize scenario-dependency by asking modellers to keep crop management fixed at present-day level or adjust it only in response to climate without any regard to the time horizons associated with adaptation or economic processes. Four models are calibrated to match present-day yield levels while LPJ-GUESS simulates potential yields assuming optimal management. Only two of the crop models allow for an adjustment of planting dates in response to climate change (GEPIC and PEGASUS, see Table 1). Three of the models keep the total heat unit sum to reach maturity constant, assuming no change in crop cultivar which effectively leads to a shortening of the growing season. Representation of soil nutrient limitation varies substantially between models, with two models (LPJ-GUESS and LPJmL) considering no soil nutrient limitation at all, while the nutrients considered and the assumptions on fertilizer application differ between the other three models. The effects of these assumptions on yield changes simulated by the different crop models are not studied here since the focus of this study is on developing efficient emulators, but these assumptions inform both the simulated yield changes as well as the emulators which attempt to imitate the behaviour of the crop models. The results of the ISIMIP crop models have been studied in detail in Rosezweig et al. (2014).

We have tested three different approaches for emulating crop yield changes simulated by GGCMs, two of which include pCO2 changes simulated by five GGCMs driven by HadGEM2-ES climate projections for four RCPs. All approaches rely on GMT as the main predictor of yield change at the grid scale. Two of the approaches include pCO2 as an additional predictor. An approach (a) attributing the yield variation within an individual GMT bin of a simulation with varying pCO2 to the change in pCO2 shows the poorest overall performance. An approach (b) based on the difference between runs with and without direct CO2 fertilization effects performs similarly well as a simple approach (c) using only GMT as a single predictor. For local (grid level) crop yields, Considering the added complexity in approach (b) compared to (c), the simple approach (c) performs slightly better than approach (b) for the LPJmL. On the other hand, appears in general preferable even though it may not provide the best result in all regions. While our tests indicate that the emulators perform better for some crop models than for yield changes weighted by actual growing areas and irrigation patterns and aggregated over large regions (i.e., regional production), approach (b) slightly outperforms approach others we strongly advise against relying solely on results from any one particular model, but instead to always consider the full range of uncertainty spanned by the GGCMs. Similarly, different GGCMs still account for more than 15% of the total variance of the ISIMIP ensemble at GMT=2.5 °C in a number regions (Figure 4) which is why emulators should be constructed for all GGCMs.

Given the availability of crop model simulations in the ISIMIP archive, emulators based on approach (a) and (c) in reproducing changes under low-warming RCPs. Considering the added complexity could be constructed for all five GGCMs for the remaining four GGCMs (IPSL-CM5A-LR, MIROC-ESM-CHEM, GFDL-ESM2M, NorESM1-M). Emulators based on approach (b) compared to (c), the simple could only be constructed for LPJmL and produced (and PEGASUS if using only RCP8.5 for training). With its five GGCMs, which were selected from the CMIP5 ensemble based primarily on data availability at the time, the ISIMIP subset likely underestimates the total uncertainty in future climate impacts attributable to GGCMs for many regions, however, the ISIMIP subset es-
sentially samples as much uncertainty as is possible with only 5 GCMs (McSweeney and Jones, 2016). The generally good performance of approach (c) appears in general preferable. This suggests that simplified predictions of large-scale agricultural yields may not require additional crop model simulations with CO$_2$ levels held at a historical level if planning to extend the GCM coverage.

The impact model ensemble available with ISIMIP data assembled in this study also indicates that the variability of crop yields is projected to increase in conjunction with increasing ΔGMT in many important regions for the four major staple crops. Such a hike in yield volatility could have significant policy implications by affecting food prices and supplies, although management assumptions as well as model-structural limitations of the GGCMs to account for crop stress factors may impact the models’ ability to project future changes in variability.

The scalability of each component (mean yields and yield variability) mean yields is conducive to the development of predictor functions relating ΔGMT, or other aggregate climate variables readily available from simplified climate models (such as pCO$_2$) to regional or global mean crop yield impacts. This lays the groundwork for a further exploration of the economic impacts of climate change encountered at target warming levels or over policy relevant regions.

Data availability. The coefficients estimated with Equations 1 to 3 are available as a Supplement, along with supplementary figures and RMSE estimates, at https://cloud.pik-potsdam.de/index.php/s/5J8vDoQvycH2muZ. The GGCM simulations that the analysis in this paper is based on are available through https://esg.pik-potsdam.de/search/isimip-ft/, with additional documentation available on the ISIMIP website https://www.isimip.org/outputdata/caveats-fast-track/.

Acknowledgements. This work was supported within the framework of the Leibniz Competition (SAW-2013 PIK-5) and by the EU FP7 HELIX project (grant no. 603864), and by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (16_II_148_Global_A_IMPACT).

For their roles in producing, coordinating, and making available the ISIMIP model output, we acknowledge the modelling groups (listed in Table 1 of this paper) and the ISIMIP cross-sectoral science team.


References


IFA: Fertilizer Use by Crop, 5th edn, International Fertilizer Industry Association (IFA), International Fertilizer Development Center (IFDC), International Potash Institute (IPI), Potash and Phosphate Institute (PPI), and Food and Agriculture Organization (FAO), Rome, 2002.


