Reprint 2017-1

Statistical emulators of maize, rice, soybean and wheat yields from global gridded crop models

Élodie Blanc

Reprinted with permission from *Agricultural and Forest Meteorology*, 236, 145–161.

The **MIT Joint Program on the Science and Policy of Global Change** combines cutting-edge scientific research with independent policy analysis to provide a solid foundation for the public and private decisions needed to mitigate and adapt to unavoidable global environmental changes. Being data-driven, the Joint Program uses extensive Earth system and economic data and models to produce quantitative analysis and predictions of the risks of climate change and the challenges of limiting human influence on the environment—essential knowledge for the international dialogue toward a global response to climate change.

To this end, the Joint Program brings together an interdisciplinary group from two established MIT research centers: the **Center for Global Change Science (CGCS)** and the **Center for Energy and Environmental Policy Research (CEEPR)**. These two centers—along with collaborators from the Marine Biology Laboratory (MBL) at Woods Hole and short- and long-term visitors—provide the united vision needed to solve global challenges.

At the heart of much of the program's work lies MIT's Integrated Global System Model. Through this integrated model, the program seeks to discover new interactions among natural and human climate system components; objectively assess uncertainty in economic and climate projections; critically and quantitatively analyze environmental management and policy proposals; understand complex connections among the many forces that will shape our future; and improve methods to model, monitor and verify greenhouse gas emissions and climatic impacts.

This reprint is intended to communicate research results and improve public understanding of global environment and energy challenges, thereby contributing to informed debate about climate change and the economic and social implications of policy alternatives.

—Ronald G. Prinn and John M. Reilly,
Joint Program Co-Directors

MIT Joint Program on the Science and Policy of Global Change

Massachusetts Institute of Technology
77 Massachusetts Ave., E19-411
Cambridge MA 02139-4307 (USA)

T (617) 253-7492 **F** (617) 253-9845
globalchange@mit.edu
http://globalchange.mit.edu

Agricultural and Forest Meteorology 236 (2017) 145–161

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Statistical emulators of maize, rice, soybean and wheat yields from global gridded crop models

Élodie Blanc

Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA

ARTICLE INFO

Article history:
Received 14 April 2016
Received in revised form
15 December 2016
Accepted 26 December 2016

Keywords:
Crop yields
Crop model
Statistical model
Climate change

ABSTRACT

This study provides statistical emulators of crop yields based on global gridded crop model simulations from the Inter-Sectoral Impact Model Intercomparison Project Fast Track project. The ensemble of simulations is used to build a panel of annual crop yields from five crop models and corresponding monthly summer weather variables for over a century at the grid cell level globally. This dataset is then used to estimate, for each crop and gridded crop model, the statistical relationship between yields, temperature, precipitation and carbon dioxide. This study considers a new functional form to better capture the non-linear response of yields to weather, especially for extreme temperature and precipitation events, and now accounts for the effect of soil type. In- and out-of-sample validations show that the statistical emulators are able to replicate spatial patterns of yields crop levels and changes overtime projected by crop models reasonably well, although the accuracy of the emulators varies by model and by region. This study therefore provides a reliable and accessible alternative to global gridded crop yield models. By emulating crop yields for several models using parsimonious equations, the tools provide a computationally efficient method to account for uncertainty in climate change impact assessments.

1. Introduction

The vulnerability of crops to weather is well known and numerous studies have attempted to estimate the impact of climate change on yields (Challinor et al., 2014). These studies generally rely on either process-based crop models (e.g. Alexandrov and Hoogenboom, 2000; Butt et al., 2005; Deryng et al., 2014; Parry et al., 1999; Rosenzweig and Parry, 1994) or statistical techniques (e.g. Blanc, 2012; Blanc and Strobl, 2013; Haim et al., 2007; Lobell and Field, 2007; Schlenker and Roberts, 2009). While process-based crop models are able to capture the effect of weather and other environmental conditions on crop growth and yields at the grid cell or site level, they are computationally demanding and sometimes proprietary, which limits their accessibility. On the other hand, statistical models are more easily applicable but depend on the availability of observations to estimate the impact of average weather conditions on crop yields while controlling for other factors. To benefit from the capabilities of processed-based models while preserving the application simplicity of statistical models, Blanc and Sultan (2015) provide an ensemble of statistical tools emulating maize yields from process-based crop models at the grid cell level globally using a simple set of weather variables. They employ the 'perfect model' approach, consisting of training a statistical model on the output of a process-based crop model, based on the assumption that these output are 'true'. This method has been used by Holzkämper et al. (2012) and Lobell and Burke (2010) with the purpose of evaluating the ability of statistical models to predict crop yields out-of-sample. These studies find that statistical models are capable of replicating the out-of-sample outcomes of process-based crop models reasonably well. Oyebamiji et al. (2015) expand on these studies by estimating a crop yield emulator at the global level for five different crops but, as in previous studies, only consider one process-based crop model. As the choice of crop model is an important source of uncertainty in climate change impact assessments on crop yields (e.g. Bassu et al., 2014; Mearns et al., 1999), Blanc and Sultan (2015) expanded the scope and applicability of statistical emulators by considering five different crop models. These emulators are based on simulations data from the Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP) Fast Track experiment dataset of global gridded crop models (GGCM) simulations. This project, coordinated by the Agricultural Model Intercomparison and Improvement Project (AgMIP) (Rosenzweig et al., 2013) as part of ISI-MIP (Warszawski et al., 2014), was tailored specifically to compare crop models. Therefore, all GGCMs simulations were driven by bias-corrected climate change projections

E-mail address: eblanc@mit.edu

http://dx.doi.org/10.1016/j.agrformet.2016.12.022

derived from the Coupled Model Intercomparison Project, phase 5 (CMIP5) archive (Hempel et al., 2013; Taylor et al., 2012). The statistical emulators produced by Blanc and Sultan (2015) provide an accessible tool to estimate the impact of climate change on crop yields while accounting for crop modeling uncertainty by allowing users to emulate yields projections from five different GGCMs. However, the crop yield emulators from Blanc and Sultan (2015) are only available for maize. This study proposes to expand the scope of these emulators to three additional crops: rice, soybean and wheat.

This study also improves the response functions estimated by Blanc and Sultan (2015) by estimating more precisely the response of crop yields to weather. The effect of weather on crop yields is non-linear and is therefore usually modeled in regression analyses by including a quadratic term in the specification (e.g. Blanc, 2012; Grassini et al., 2013; Schlenker and Lobell, 2010). However, the symmetrical concave relationship imposed by this functional form might be too restrictive. Blanc and Sultan (2015) find that a fifth order polynomial transformation is well suited to represent the nonlinear relationship between weather and crop yields. However, the polynomial form exhibits behaviors difficult to explain for extreme values of temperature and precipitation. As an alternative, this study applies the fractional polynomial method from Royston and Altman (1994). This approach provides the flexibility and improved fit of a non-parametric model, but with the simplicity of a parametric model.

Data and methods used to statistically estimate relationship between yields and weather variables are presented in Section 2. Results are presented and discussed in Section 3. The models are validated in Section 4. Section 5 concludes.

2. Material and methods

2.1. Data

Data used in this study are sourced from the ISI-MIP Fast Track experiment, an inter-comparison exercise of global gridded process-based crop models using the CMIP5 climate simulations.[1] In this exercise, several modeling groups provided results from global gridded process-based crop models run under the same set of weather and CO_2 concentration inputs.

2.1.1. Weather and CO_2

Bias-corrected weather data used as input into each crop model are obtained from the CMIP5 climate data simulations. Daily weather data generated by three CMIP5 climate models, or General Circulation Models (GCMs): HadGEM2-ES, NorESM1-M, and GFDL-ESM2 M. These GCMs are selected to be representative of respectively, high, medium and low levels of global warming (Warszawski et al., 2014).

GCM simulations are provided for the 'historical' period of 1975–2005 and the 'future' period of 2006–2099. For the 'future' period, one Representative Concentration Pathway (RCP) consistent with the highest level of global warming compared to historical conditions, RCP 8.5, and the corresponding CO_2 concentrations data (Riahi et al., 2007) is considered.[2] Combined with the large range of climate change patterns represented by the three GCMs, this study considers the broadest plausible range of future climate change.

Using daily precipitation, and minimum and maximum temperature produced by each GCM and used as inputs by GGCMs, monthly averages of precipitation (*Pr*) and temperature (*Tmean*) are calculated for each summer month.[3] For ease of reference, in this study numbers suffixes are used to represent each summer month, so _1, _2, and _3 refer to, respectively, June, July and August in the Northern Hemisphere and December, January and February in the Southern Hemisphere.

2.1.2. Crop yields

Crop yields are obtained from GGCMs members of the ISI-MIP Fast Track experiment. Due to data limitations, simulations from five crop models are selected: the Geographic Information System (GIS)-based Environmental Policy Integrated Climate (GEPIC) model (Liu et al., 2007; Williams, 1995), the Lund Potsdam-Jena managed Land (LPJmL) dynamic global vegetation and water balance model (Bondeau et al., 2007; Waha et al., 2012), the Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS) with managed land model (Bondeau et al., 2007; Lindeskog et al., 2013; Smith et al., 2001), the parallel Decision Support System for Agrotechnology Transfer (pDSSAT) model (Elliott et al., 2013; Jones et al., 2003), and the Predicting Ecosystem Goods And Services Using Scenarios (PEGASUS) model (Deryng et al., 2011). For each of these GGCMs, model simulations considering the effect of CO_2 concentrations are selected in order to account for the CO_2 fertilization effect, which plays an important role on biomass production. In this study, only simulations assuming no irrigation are considered in order to capture the effect of precipitation on crop yields.

All GGCMs estimate annual crop yields in metric tons per hectare (t/ha) at a $0.5 \times 0.5^{\circ}$ resolution (about 50 km^2). And although they differ in their representation of crop phenology, leaf area development, yield formation, root expansion and nutrient assimilation, they all account for the effect of water, heat stress and CO_2 fertilization, and assume no technological change. A more detailed description of each model's processes is provided by Rosenzweig et al. (2014). As mentioned in Blanc and Sultan (2015), caveats are associated with each model leading to divergences and GGCM-specific periodic patterns of yield projections.[4]

Crop models simulate yields from 1975 to 2005 for the 'historical' period and 2006–2099 for the 'future' period. As only one RCP scenario is selected for each GCM, the panel is constructed over the consecutive period 1975–2099 without distinction (i.e. one historical scenario and one future scenario for each GCM). In the final sample, grid cells for which there are less than 10 yield observations after data cleaning are omitted.

2.1.3. Soil orders

Soil orders at the $0.5 \times 0.5^{\circ}$ resolution are extracted from the FAO-UNESCO (2005) Soil Map of the World using the USDA soil taxonomy (Soil Survey Staff, 1999) which classifies soils on the basis of soil physical and chemical properties observed in situ (e.g. soil horizons,[5] structure, texture, color) and inferred from environmental conditions (e.g., soil temperature and moisture regimes). As shown in Fig. 1, soils are grouped into 12 main soil orders: Gelisols which are permanently frozen or contain permafrost near the soil surface and are found in the Arctic, Antarctic and extremely high elevations; Histosols which are composed mainly of decomposing organic matters and forming in areas of poor drainage; Spodosols which develop under coniferous vegetation where litter contributes to acid accumulations in the soil; Andisols which

[1] The data are available for download at https://www.pik-potsdam.de/research/climate-impacts-and-vulnerabilities/research/rd2-cross-cutting-activities/isi-mip/data-archive/fast-track-data-archive

[2] The data are available at http://tntcat.iiasa.ac.at/RcpDb/dsd?Action=htmlpage&page=welcome

[3] Mean temperature is calculated as $Tmean = (Tmin + Tmax)/2$, Tmin and Tmax are, respectively, the minimum and maximum daily temperatures.

[4] These caveats are discussed at https://www.pik-potsdam.de/research/climate-impacts-and-vulnerabilities/research/rd2-cross-cutting-activities/isi-mip/data-archive/fast-track-data-archive/data-caveats

[5] Soil order data are available for download at https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/use/?cid=nrcs142p2_054013

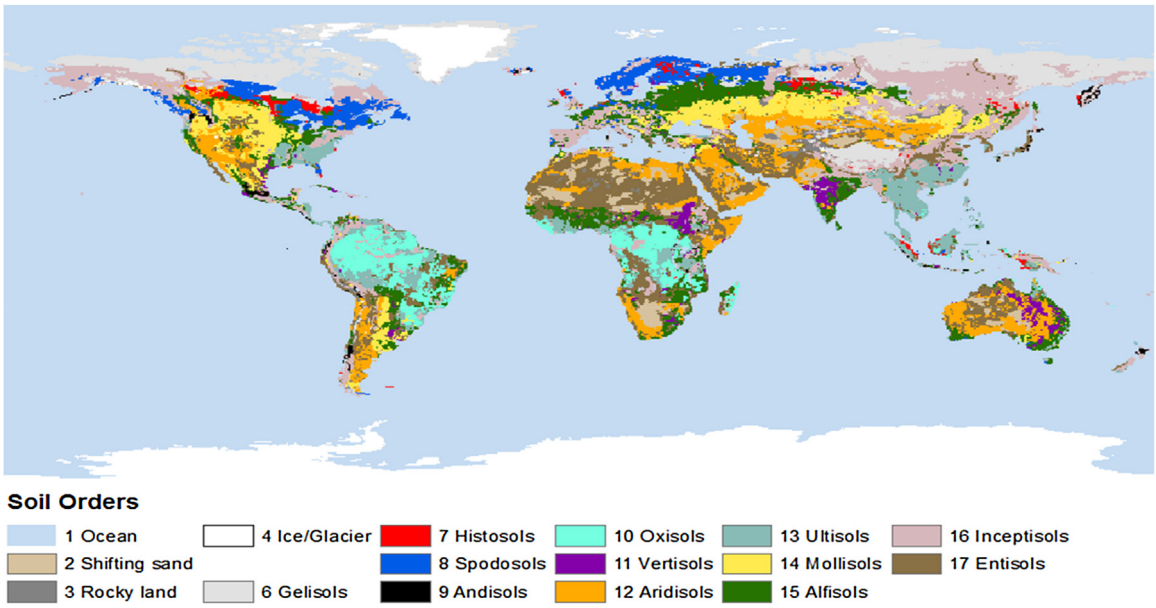

Fig. 1. Global soil regions based on the FAO-UNESCO Soil Map of the World using the USDA soil taxonomy.

develop from volcanic materials; Oxisols which are found in tropical and subtropical latitudes where precipitation and temperature are high; Vertisols which are clay-rich exhibiting seasonal cracking and swelling in climates with a distinct dry seasons; Aridisols which develop in dry environments and are prone to salinization; Ultisols which are red clay soils formed in humid areas that are intensely leached; Mollisols which are dark, rich and fertile soils commonly used for pastures and crops; Alfisols which are formed under forest vegetation in semiarid to humid areas; Inceptisols which are young soils occurring in a wide range of climates except aridic; and Entisols which are mineral soils unaltered from their parent's sediment or rock material.

2.1.4. Summary information and statistics

The size characteristics of the panel dataset are summarized in Table 1. Samples have on average 18 million observations covering about 60,000 grid cells globally. However, sample sizes vary by crop and GGCM, with simulations for maize and wheat being the most extensive. Simulations from pDSSAT for rice and soybean, and from PEGASUS for rice are not available. Additionally, simulations for wheat by the pDSSAT model are only available for the HadGEM2 GCM, hence the sample size for this crop model when predicting wheat yields is smaller than those for other crop models.

Table 1
GGCMs summary information.

Crop	Model	Observations	Grid Cells
Maize	GEPIC	22,293,247	62,005
	LPJ-GUESS	20,665,195	56,620
	LPJmL	22,794,487	62,148
	PEGASUS	13,406,155	51,580
	pDSSAT	15,758,066	51,447
Rice	GEPIC	22,356,067	62,249
	LPJ-GUESS	19,638,299	55,834
	LPJmL	21,874,607	59,169
Soybean	GEPIC	22,277,853	62,115
	LPJ-GUESS	20,189,609	55,585
	LPJmL	22,469,951	61,367
	PEGASUS	9,534,992	43,436
Wheat	GEPIC	22,987,936	63,260
	LPJ-GUESS	19,816,608	55,106
	LPJmL	24,151,787	65,732
	PEGASUS	13,551,266	51,413
	pDSSAT	5,298,321	50,669

Notes: simulations for wheat from the pDSSAT model are only available for the HadGEM2 GCM.

As depicted in Fig. 2, the growing season is from May to September in the Northern Hemisphere, and from mid-November to March in the Southern Hemisphere for maize; from mid-May to October in the Northern Hemisphere and late November to mid-April in the Southern Hemisphere for rice; and from mid-May to early October in the Northern Hemisphere and late November to March in the Southern Hemisphere for soybean. For wheat, which represents a combination of spring and winter wheat, two main planting seasons are discerned: spring and fall. Maturity largely occurs at the end of summer in the Northern hemisphere but is more dispersed in the Southern hemisphere. See Figure A1 in Appendix A for detailed representation of the distribution of planting and maturity date for each crop and hemisphere.

Summary statistics for crop yields by GGCM and GCM are presented in Table 2 Global average crop yields are generally the smallest for the LPJmL model and the highest for the PEGASUS model—although this model provides simulations for maize and wheat only. The range of simulated yields vary greatly across models, with the pDSSAT model simulating maize yields of up to 34.9t/ha compared to a maximum of 9.7t/ha projected by the LPJ-GUESS model. For rice, soybean and wheat, the smallest maximum yields are projected by the GEPIC model with 11.3t/ha, 5.8t/ha and 10.4t/ha respectively, while the upper bound is projected by the LPJmL for rice (19.5t/ha), PEGASUS for soybean (18.3t/ha) and pDSSAT for wheat (34.3t/ha). Across GCMs, crop yields are on average the largest under the NorESM1_M scenario although the range of projected yields vary greatly depending on the GGCM and crop considered.

Summary statistics for the summer weather variables, *Tmean and Pr*, and *CO2* are presented in Table 3. For clarity purposes, summary statistics for these variables are averaged over all GGCMs (weather inputs differ slightly by crop model due to different spatial coverage, i.e., a different number of grid cells are represented by each GGCM for each crop). Summary statistics detailed by GGCM are provided in Appendix A. Precipitation is on average the lowest in the first month of summer and the highest in the last month. Temperatures, however peak in the second month of summer.

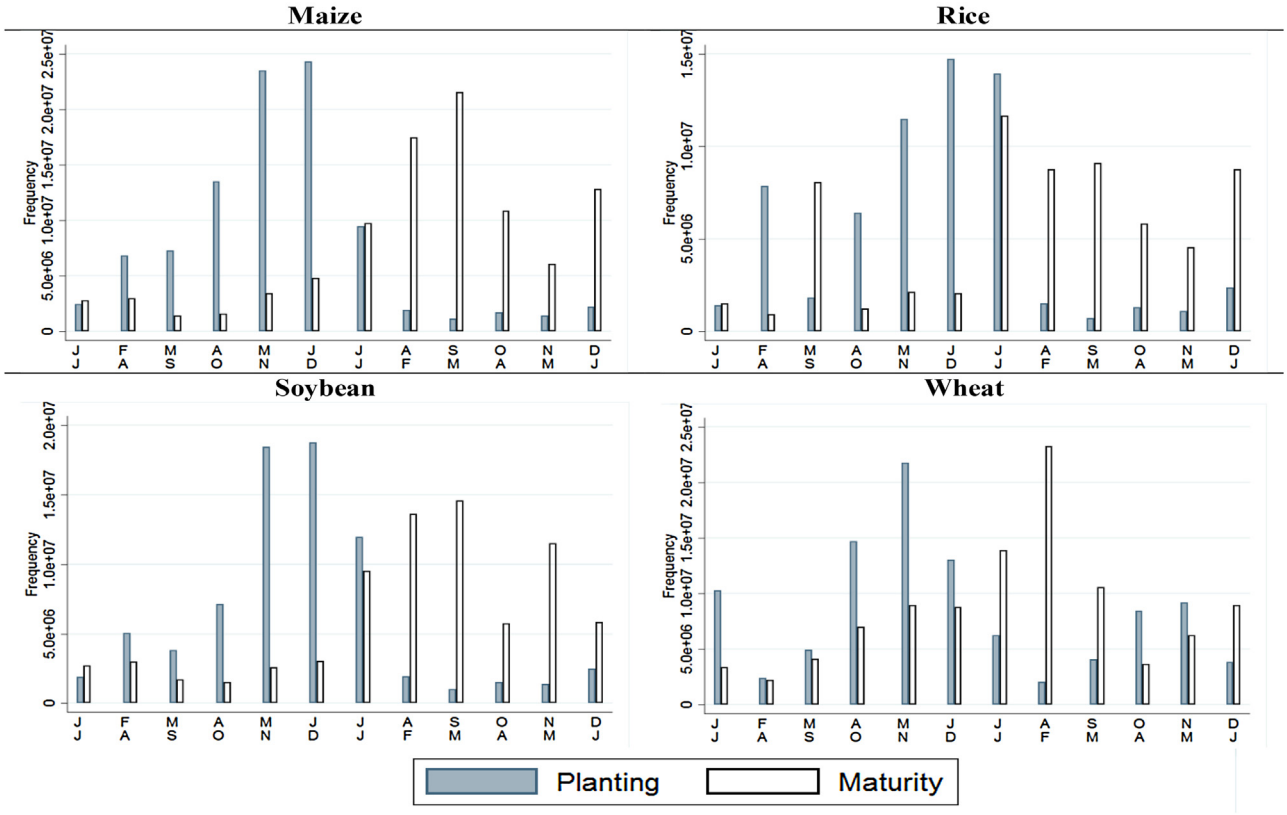

Fig. 2. Monthly distribution of planting and maturity dates by crop.
Note: Planting and maturity dates in the Southern hemisphere (on the second line of the axis) are rescaled in order for seasons to coincide with those in the Northern hemisphere.

Table 2
Summary statistics for crop yields (t/Ha) by GGCM and GCM.

Crop	Model	GFDL_ESM2M			HadGEM2_ES			NorESM1_M		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Maize	GEPIC	1.8	0.0	14.7	1.6	0.0	12.3	1.9	0.0	12.8
	LPJ-GUESS	1.7	0.0	10.3	1.8	0.0	10.8	1.9	0.0	9.7
	LPJmL	1.3	0.0	17.4	1.5	0.0	17.7	1.5	0.0	17.2
	pDSSAT	2.6	0.0	24.1	2.9	0.0	23.9	2.9	0.0	23.8
	PEGASUS	1.8	0.0	34.6	1.7	0.0	34.4	2.0	0.0	34.9
Rice	GEPIC	1.6	0.0	13.6	1.5	0.0	11.3	1.7	0.0	11.3
	LPJ-GUESS	1.1	0.0	16.3	1.1	0.0	18.6	1.3	0.0	17.2
	LPJmL	1.1	0.0	17.1	1.1	0.0	19.5	1.2	0.0	17.8
Soybean	GEPIC	0.9	0.0	5.8	0.9	0.0	5.0	1.0	0.0	5.8
	LPJ-GUESS	0.8	0.0	8.2	0.8	0.0	8.9	1.0	0.0	10.3
	LPJmL	0.7	0.0	13.7	0.8	0.0	13.1	0.9	0.0	15.2
	PEGASUS	1.2	0.0	18.0	1.1	0.0	16.5	1.4	0.0	18.3
Wheat	GEPIC	1.3	0.0	10.4	1.4	0.0	11.0	1.4	0.0	10.8
	LPJ-GUESS	2.3	0.0	17.1	2.3	0.0	15.0	2.4	0.0	14.4
	LPJmL	1.5	0.0	17.0	1.5	0.0	16.3	1.5	0.0	16.2
	pDSSAT				2.1	0.0	34.3			
	PEGASUS	1.2	0.0	25.8	1.0	0.0	26.1	1.2	0.0	27.9

Table 3
Summary statistics for averaged weather variables by GCM at the global level.

Variable	Unit	GFDL_ESM2M			HadGEM2_ES			NorESM1_M		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Pr_1	mm/day	3.2	0.0	147.1	3.0	0.0	152.1	3.0	0.0	148.6
Pr_2	mm/day	3.5	0.0	176.0	3.5	0.0	173.6	3.5	0.0	189.0
Pr_3	mm/day	3.5	0.0	127.3	3.5	0.0	112.5	3.5	0.0	102.3
Tmean_1	°C	21.4	−3.1	45.0	22.8	−3.1	46.6	22.0	−3.3	43.6
Tmean_2	°C	23.1	0.5	45.1	24.5	0.6	47.0	23.9	0.0	44.8
Tmean_3	°C	22.4	−1.2	45.5	23.8	−1.3	46.6	22.9	−2.1	44.7

Note: suffixes _1, _2, _3 denote, respectively, June, July and August in the Northern Hemisphere and December January and February in the Southern Hemisphere.

Table 4
Summary statistics for averaged weather variables averaged over GCMs at the soil order level.

Subsample	Tmean_1			*Tmean_2*			*Tmean_3*			*Pr_1*			*Pr_2*			*Pr_3*		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
7 Histosols	16.8	−8.74	37.1	20.1	1.29	39.4	19.9	2.86	39.4	3.00	0.0	59.7	3.38	0.0	48.2	3.30	0.0	48.2
8 Spodosols	13.9	−10.2	36.6	17.7	0.17	38.7	17.6	0.83	38.7	2.41	0.0	57.4	2.66	0.0	35.7	2.68	0.0	30.7
9 Andisols	17.5	−2.62	35.7	19.3	0.29	35.0	19.8	1.71	35.1	3.58	0.0	114	3.98	0.0	114	4.17	0.0	90.8
10 Oxisols	27.4	10.1	40.7	27.1	11.5	39.8	27.2	13.2	41.5	7.57	0.0	87.6	7.84	0.0	91.8	7.91	0.0	112
11 Vertisols	30.0	5.14	43.7	29.6	9.42	44.8	29.1	11.1	45.8	2.79	0.0	84.5	4.49	0.0	105	4.96	0.0	105
12 Aridisols	25.6	−0.82	46.3	27.8	1.76	47.3	27.7	2.72	47.3	0.79	0.0	55.4	0.99	0.0	55.4	1.09	0.0	109
13 Ultisols	27.0	−1.87	42.1	27.5	2.71	42.1	27.6	9.15	41.8	6.87	0.0	156	7.41	0.0	189	7.41	0.0	189
14 Mollisols	20.4	−2.31	41.5	23.5	3.76	42.0	23.5	3.94	42.0	2.10	0.0	80.7	2.22	0.0	114	2.00	0.0	114
15 Alfisols	23.0	−2.95	44.7	24.5	2.76	45.2	24.3	4.70	46.1	3.05	0.0	145	3.79	0.0	158	4.01	0.0	158
16 Inceptisols	16.9	−10.7	44.5	20.0	−1.92	44.4	19.9	0.62	43.8	3.04	0.0	116	3.68	0.0	135	3.75	0.0	135
17 Entisols	27.6	−9.29	46.4	28.7	−2.29	46.4	28.6	1.80	46.7	2.20	0.0	117	2.86	0.0	119	3.17	0.0	119

While no clear pattern amongst GCMs is discernable from these statistics in terms of precipitation, temperatures are clearly the highest under the HadGEM2-ES GCM and the lowest under the GFDL-ESM2M GCM. Atmospheric CO_2 concentrations, not represented in this table, do not differ by GCM or GGCM and range from 331 parts per million (ppm) in 1975 to 927 ppm by the end of the century.

Weather statistics for each soil order averaged over GCMs are reported in Table 4. These statistics indicate that climates differ across soil regions with mid-summer temperature averaging between 18° C in regions with Spodosols soils to 30° C in regions with Vertisols soils. Similarly, precipitation ranges from less than 1 mm/day in regions with Aridisols soils to more than 7 mm/day in regions with Oxisols soils.

2.2. Methods

A statistical model is fitted for each crop to a panel of yields produced by a GGCM. The response functions are then used to predict crop yields. To evaluate the accuracy of the emulator, yield projections from the emulator are then compared to the outcome of the process-based crop models using the same weather inputs—in-sample validation—and using weather from alternative climate change scenarios—out-of-sample validation. The goal of the study is to produce simple equations that emulate crop yields and can be used by others to predict changes in yields based on data from alternative GCMs. By providing emulators for an ensemble GGCMs, these emulators also allow users to account for crop modeling uncertainty in climate change impact assessments.

To statistically estimate the determinants of crop yields, Blanc and Sultan (2015) considered the effect on maize yields of a large set of climate variables, such as minimum and maximum temperatures, growing degree days and the number of days with extreme low and high temperatures, as well as the number of days without precipitation. The results showed only marginal improvements in the projection accuracy of the emulators relative to regressions that only include average summer precipitation and temperature variables (S1 specifications). In this study, the effect of within-day differences in temperature on all crops was further investigated by including the diurnal temperature range in Blanc and Sultan (2015)'s preferred specification (S1polyint). However, the results show very small improvements in the goodness of fit.[6] Therefore, this study considers a parsimonious specification that only includes average summer precipitation and temperature weather variables, as well as CO_2 and interactions between these variables. For continuity with Blanc and Sultan (2015), this specification is also labeled S1.

For maize, rice and soybean, the S1 specification includes monthly average of summer weather variables to represent a common growing season.[7] For wheat, in order to capture the effect of weather on both spring and winter wheat, the growing season covers the last month of spring and the first two months of summer.[8] Among various representations of weather effects on crop growth, this set of monthly weather variables was found to provide the best compromise in term of predictive ability and simplicity. The specification also includes interaction terms between temperature and precipitation, between temperature and CO_2 to account for the CO_2 enrichment effect, and between precipitation and CO_2 to account for the CO_2 effect on water use efficiency. For each crop and GGCM, the specification estimated is of the form:

$$\begin{aligned} Yield_{lat,lon,gcm,y} &= \alpha + \sum_{i=1}^{3} \beta_i Pr\_i_{lat,lon,gcm,y} \\ &+ \sum_{i=1}^{3} \theta_i Tmean\_i_{lat,lon,gcm,y} + \vartheta CO2_{gcm,y} + \\ &\sum_{i=1}^{3} \gamma_i Pr\_i_{lat,lon,gcm,y} * Tmean\_i_{lat,lon,gcm,y} \\ &+ \sum_{i=1}^{3} \omega_i Pr\_i_{lat,lon,gcm,y} * CO2_{gcm,y} + \\ &\sum_{i=1}^{3} \kappa_i Tmean\_i_{lat,lon,gcm,y} * CO2_{gcm,y} + \delta_{lat,lon} + \rho_{lat,lon,gcm,y} \end{aligned} \quad (1)$$

where for each year, y, *Yield* corresponds to crop yields simulated by process-based crop models for each grid cell (defined by its longitude, *lon*, and latitude, *lat*) under each climate model, *gcm*; *Pr* and *Tmean* variables correspond to monthly mean precipitation and temperature variables. *CO2* is the annual midyear CO_2 concentration level in the atmosphere; δ is a grid cell fixed effect; and ρ an error term. Following Blanc and Sultan (2015), adjustments to the specification are made for the pDSSAT model to account for soil fertility erosion, and for the GEPIC model to account for 30-yearly input of CO_2.

Eq. (1) represents a linear effect of weather and CO_2 on crop yields. However, it has been established in the literature that crop yield response to weather is non-linear. The most common and straightforward method used to represent non-linear effects in regression analyses consist of including a quadratic term for precipitation temperature, precipitation and CO_2. As detailed in Table 5, this quadratic transformation is represented by specifica-

[6] Differences in normalized root mean square error (NRMSE) when the diurnal temperature range is included and excluded were less than 0.001 percentage point.

[7] That is, the growing season for these crops is modeled as the months of June, July and August in the Northern Hemisphere, and the months of December, January and February in the Southern Hemisphere.

[8] That is, the growing season for wheat is modeled as the months of May, June and July in the Northern Hemisphere and the months of November, December and January in the Southern Hemisphere.

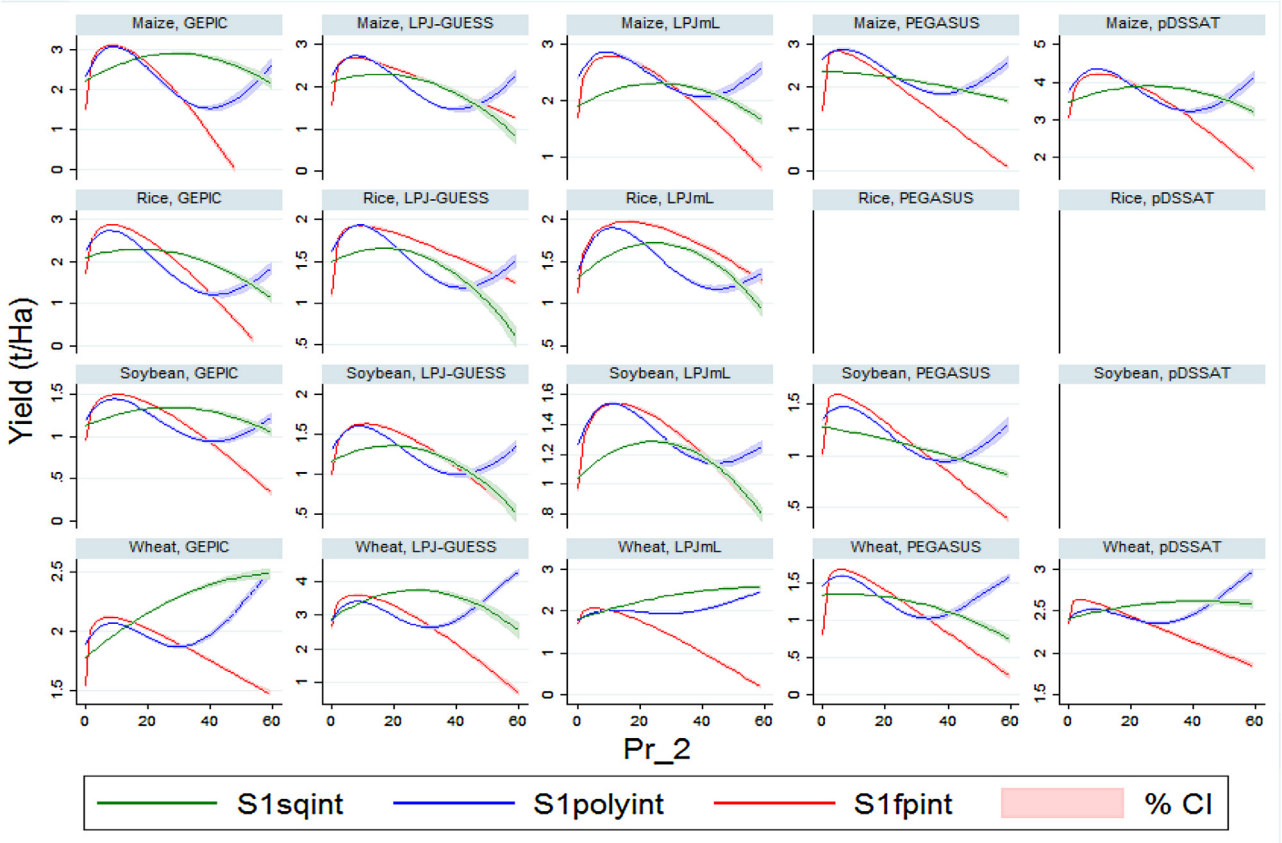

Fig. 3. Effect of Pr_2 on yields by crop and GGCM for the S1sqint, S1polyint and S1fpint specification. *Note*: covariates are held at their mean values.

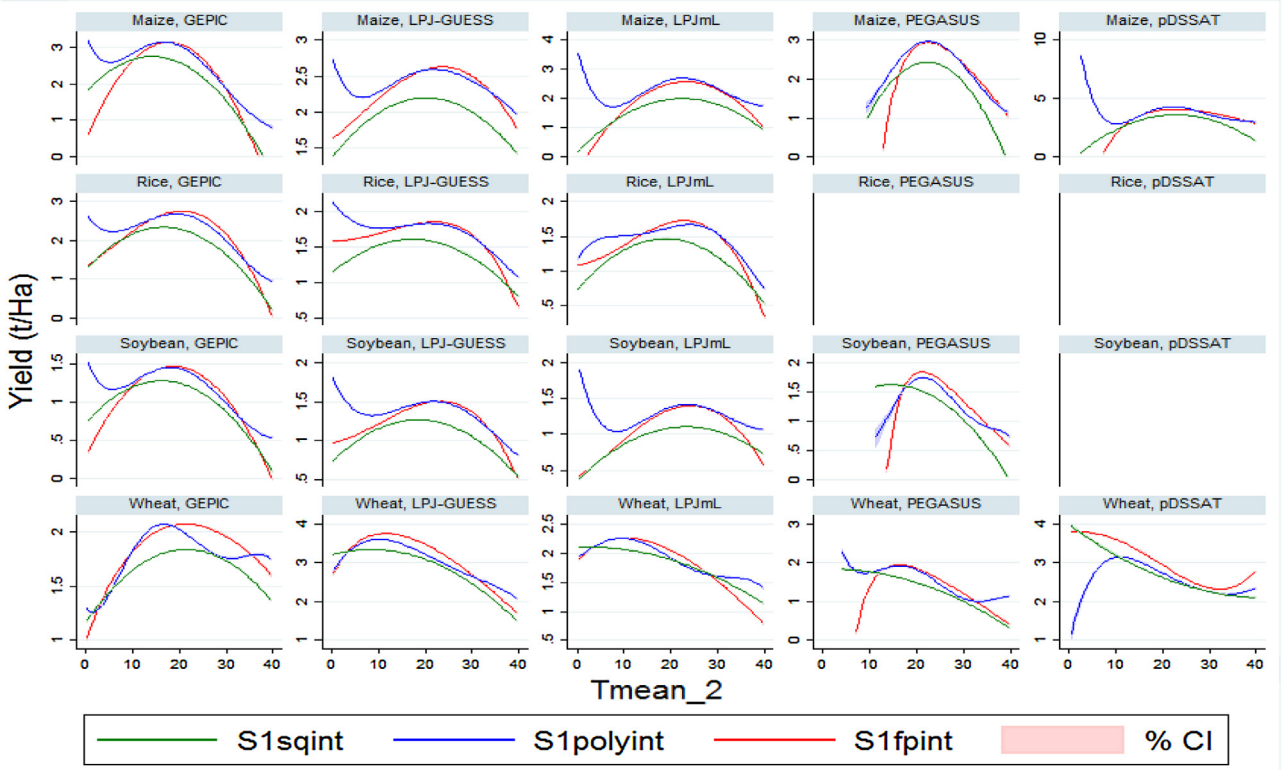

Fig. 4. Effect of Tmean_2 on yields by crop and GGCM for the S1sqint, S1polyint and S1fpint specification. *Note*: covariates are held at their mean values.

Table 5
Specification description.

Specification	Variables non-linear transformations	Sample
S1sqint	*Pr, Pr_sq, Tmean, Tmean_sq, CO2, CO2_sq*	Global
S1polyint	*Pr, Pr_sq, Pr_cu, Pr_qu, Pr_qc, Tmean, Tmean_sq, Tmean_cu, Tmean_qu, Tmean_qc, CO2, CO2_sq*	Global
S1fpint	*Pr_p1, Pr_p2, Tmean_p1, Tmean_p2, CO2_p1, CO2_p2*	Global
S1fpintsoil	*Pr_p1, Pr_p2, Tmean_p1, Tmean_p2, CO2_p1, CO2_p2*	Soil order level

Note: suffix _sq denotes square terms, _cu cubic terms, _qu quartic terms, and _qc quintic terms, _p1 and _p2 power terms.

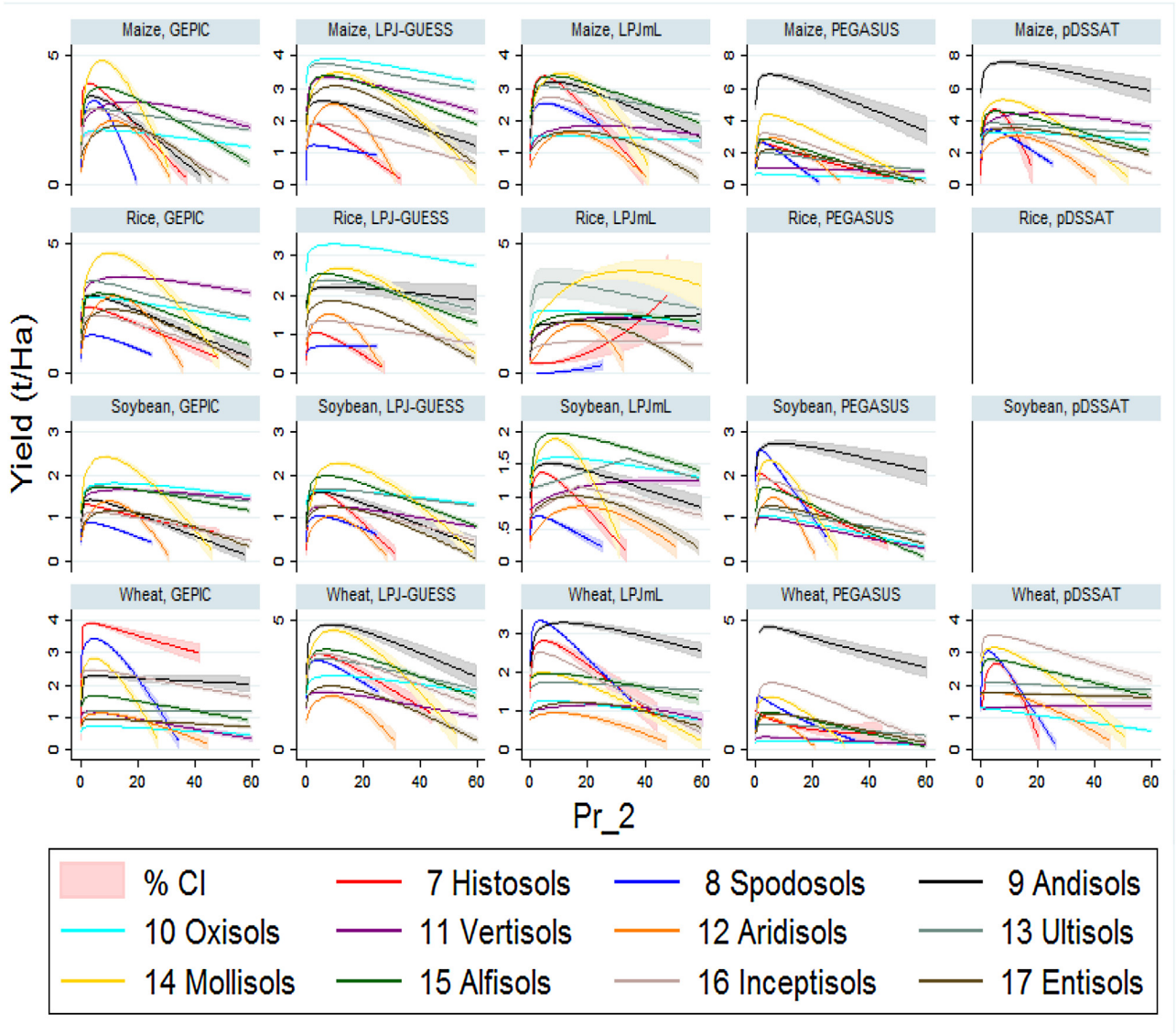

Fig. 5. Effect of Pr_2 on yields by crop and GGCM for the S1fpintsoil specification.

tion S1sqint.[9] However, this functional form imposes a constraint of symmetrical relationship. A fifth order polynomial specification, S1polyint, is thus considered to allow greater flexibility in the representation of the effect of weather. A shortcoming associated with the representation of weather effect using polynomial equation is odd tail-end behaviors and, as noted by Blanc and Sultan (2015), the weather effects should be interpreted with caution when considering extreme events. In response to this limitation, we consider a fractional polynomial specification, S1fpint, which addresses this issue by relaxing the symmetry constraint but allowing non-parametric flexibility.

In its general form, the fractional polynomial model of degree m is defined as:

$$Y = \alpha_0 + \sum_{i=1}^{m} \alpha_i X^{(p_i)} + \mu \tag{2}$$

where the parentheses on the power term on X imply the following transformation:

$$X^{(p_i)} = \begin{cases} X^{p_i} \text{ if } p_i \neq 0 \\ lnX \text{ if } p_i = 0 \end{cases} \tag{3}$$

[9] For consistency and comparison with the results of Blanc and Sultan (2015), similar specification notations are used.

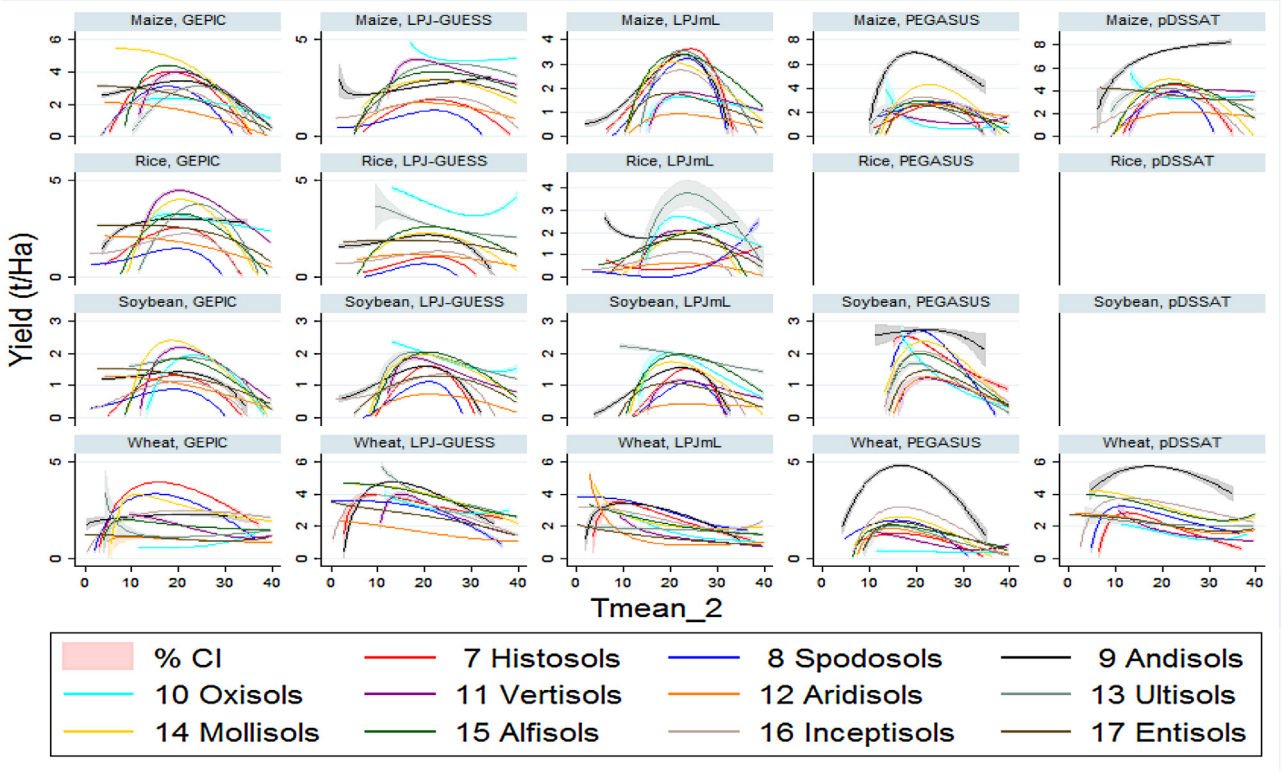

Fig. 6. Effect of Tmean_2 on yields by crop and GGCM for the S1fpintsoil specification.

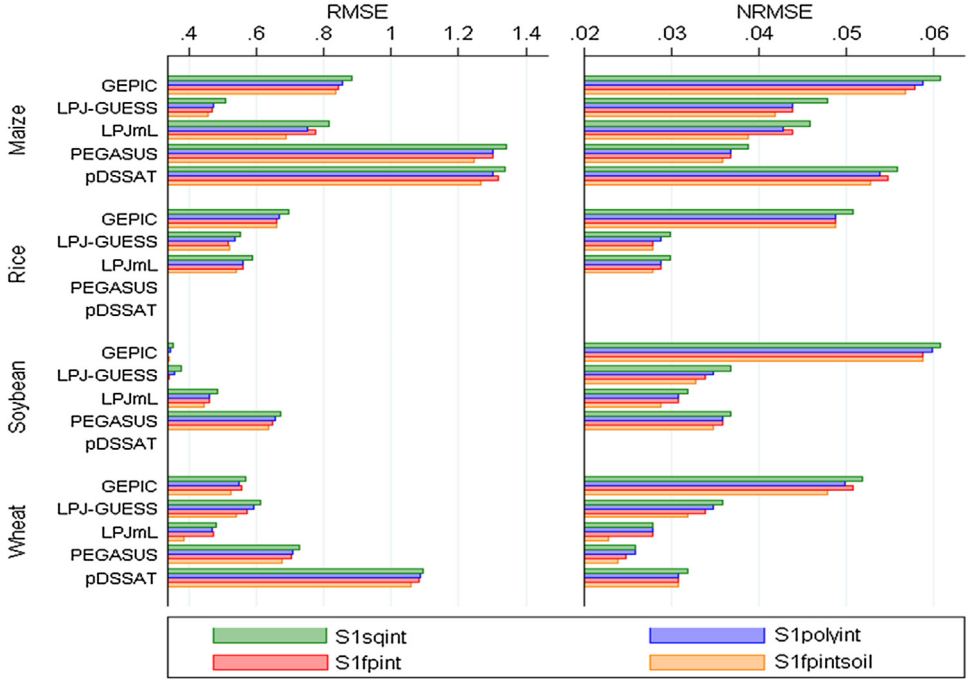

Fig. 7. Goodness of fit measures by crop and GGCM for the statistical emulators (S1sqint, S1polyint, S1fpint and S1fpintsoil specifications).

For each repeated power, p_i, the term is multiplied by another ln$X$. To fit a multivariable fractional polynomial model, and a closed-test algorithm performs a backward elimination starting from the most complex specification. In this application, the maximum permitted degree is $m = 2$. Following Royston and Sauerbrei (2008), powers are chosen from among the set $\{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$.

An important factor of crop growth and yields that has been omitted so far is soil characteristics. Depending on the soil composition, water and nutrient holding capacities differ and affect their availability to plants. Except for PEGASUS,[10] all GGCMs considered

[10] PEGASUS uses only available water capacity from the ISRIC-WISE dataset (Batjes, 2006).

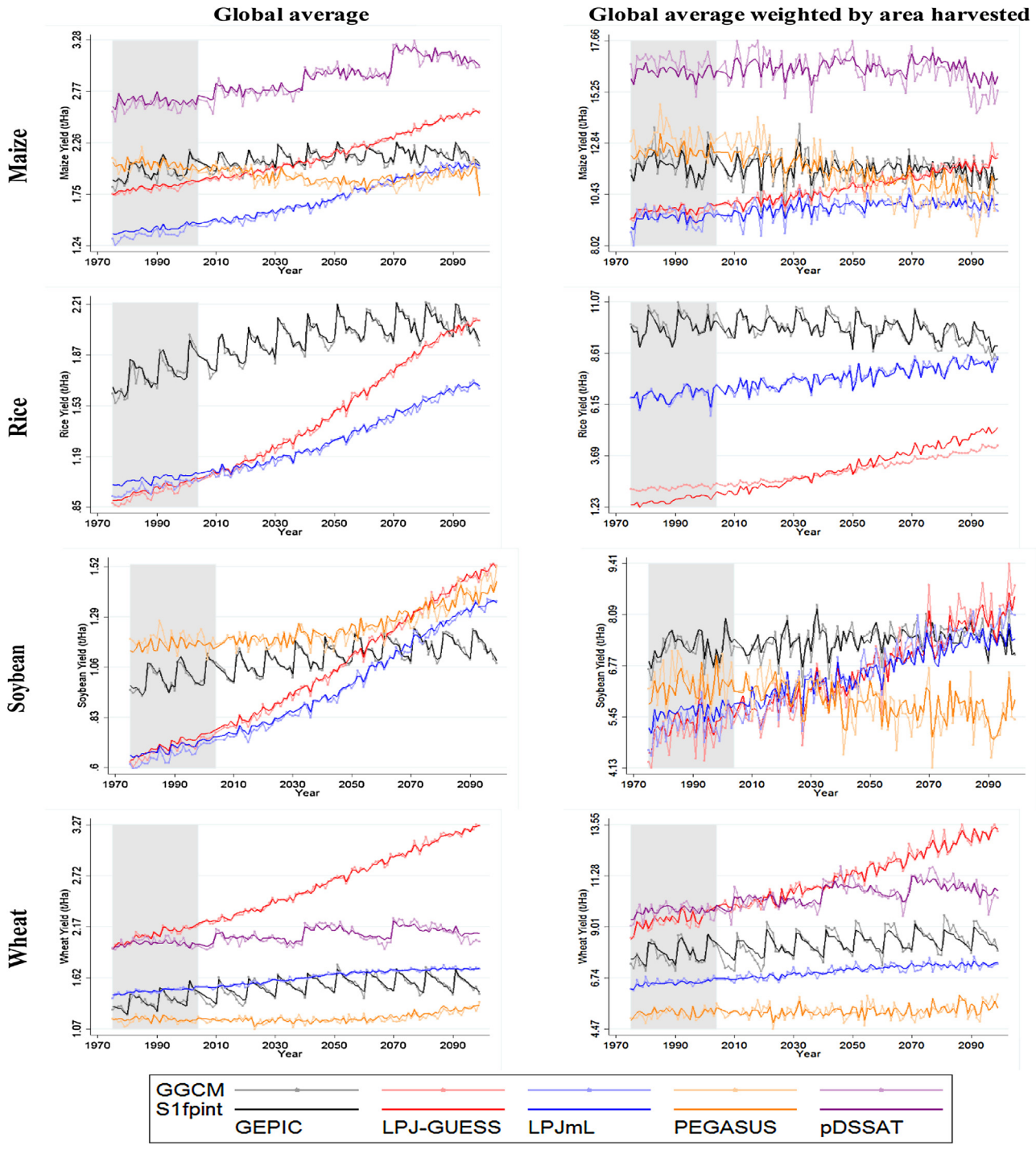

Fig. 8. Average crop yield projections from GGCMs and statistical emulators (S1fpintsoil specification). *Note*: Shaded areas represents the 'historical' period.

use soil information from the Harmonized World Soil Database, which contains more than 15,000 soil mapping units combining national and regional soil profiles with the FAO-UNESCO Soil Map of the World. To simplify their representation, general soil orders following the USDA soil taxonomy are considered.[11] As these soil types are time-invariant, their direct effects are captured by the grid-cell fixed effects. Additionally, as the weather effect on crops is expected to differ across soil types, the preferred estimation strategy estimates separate weather response functions for each soil order.[12] The specification is labeled S1fpintsoil and uses the same independent variables and functional forms as in the S1fprint specification.

[11] Trial regression estimations for the PEGASUS model using different subclasses of water capacity to capture differences in soil characteristics did not provide satisfying results. Regressions by soil orders subsamples have higher forecast performance for this model.

[12] In this analysis, we do not estimate a response function for the Gelisols soil order, as it represents soils permanently frozen.

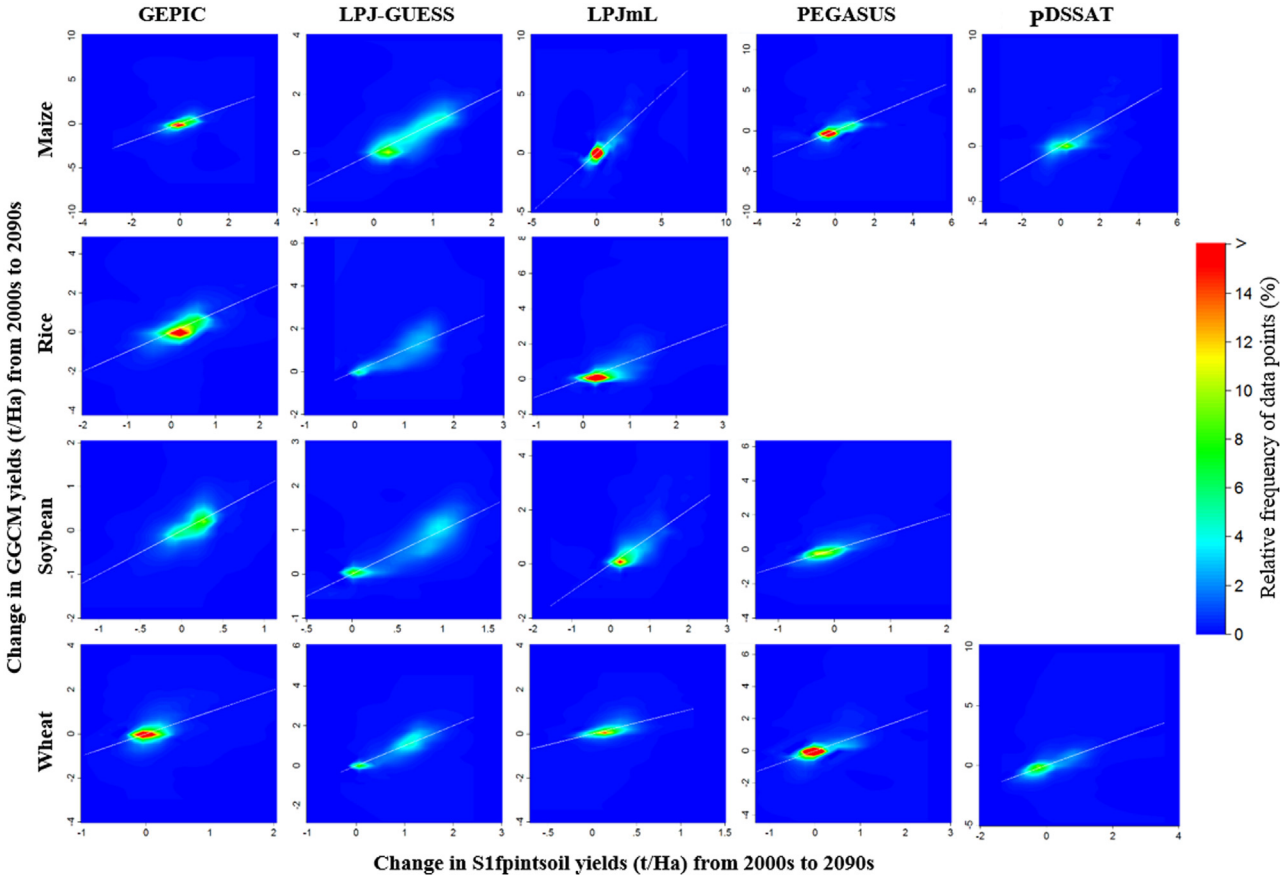

Fig. 9. Comparison of changes in crop yields from 2000 s to 2090 s between GGCMs and statistical emulators (S1fpintsoil specification).

3. Results

Regressions are first estimated for the three specifications S1sqint, S1polyint and S1fpint for each crop and GGCM at the global level. The preferred specification (S1fpint) is then re-estimated for each soil order subsample (S1fpintsoil). The power terms used for both specifications S1fpint and S1fpintsoil are reported in Appendix B. Results for each regression are presented in Appendix C. The corresponding estimated values for δ (the grid cell fixed effect) are provided in Appendix D.

At the global level, S1fpint regression results show that precipitation and temperature during all the months of the growing seasons for each crop have a significant impact on crop yields from all GGCMs under any of the three functional forms considered. The significant coefficients for the interaction terms between precipitation and temperature, *Pr_x_Tmean*, indicate that the impact of a change in temperature depends on the amount of precipitation and vice versa. However, the representation of the non-linear relationship between weather variables and yields differ between specifications. To facilitate the comparison between the three specifications, the average effect of temperature and precipitation, holding covariates at their mean values, are depicted for each crop model and GCM in Figs. 3 and 4. The graphs show that the S1sqint specification functional form, due to its symmetrical nature, is very restrictive. For instance, it shows a concave effect of precipitation on yields with a turning point at very high level of precipitation (around 30 mm/day). However, such precipitation rarely occurs in the dataset (mean *Pr_2* is around 3 mm/day). The S1polyint and S1fpint address this issue by fitting a curve skewed toward low values of precipitation. These specifications capture the skewness of the curve with a generally sharper increase in yields associated with very low precipitation than the S1sqint. However, the graphs show that the S1polyint specification represents an increase in yield for precipitation over 30 mm/day. Such tail-end behavior could lead to erroneous conclusions when extrapolating beyond the range of commonly observed precipitation, which is of particular concern when simulating the effect of climate change. The S1fpint model provides a solution to the odd tail-end behavior when using a degree-2 fractional polynomial.

The effects of temperature changes in the second summer month, again when covariates are held constant at their mean values, are shown in Fig. 4. The S1fpint model allows a better representation of the large beneficial impact of a temperature increase when temperatures are low compared to the quadratic specification S1sqint, which generally has a flatter bell shape, and the polynomial specification S1polyint, which exhibit hard to explain tail-end behavior for most GGCMs (for example, an ever-increasing positive effect of temperature on crop yields beyond 45° C). Graphs of the average effects of temperature and precipitation during the first and last months of the summer growing season, provided in Appendix E, show similar patterns.

Building on the S1fpint specification, the associated power terms are estimated for each soil order subsample in the S1fpintsoil estimations. As shown in Figs. 5 and 6, the average effect of temperature and precipitation during the second month of summer (holding covariates at their mean values within each subsample) differ greatly depending on the soil order sample considered. The largest yield response to precipitation and temperature are simulated in general in Andisols and Mollisols subsamples, which are generally representative of very fertile soils. The shape of the

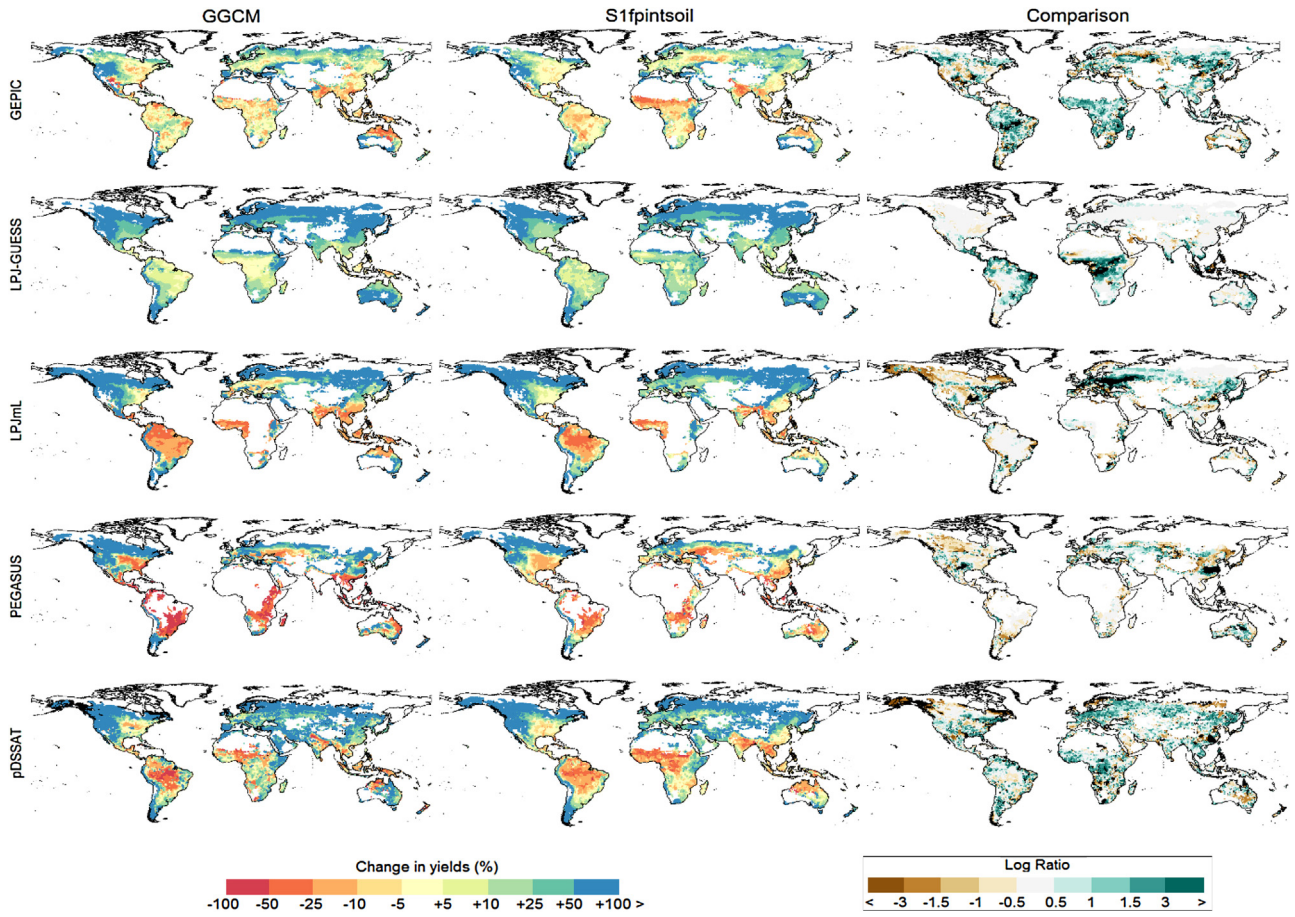

Fig. 10. Changes in maize yields from 2000 s to 2090 s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio).
Note: Grid cells where yields projections from crop models are on average less than 1t/ha over the whole study period are masked in white. Grid cells for which the sign of the impact projected with the emulator is contrary to the sign of the impact projected by the GGCM are masked in black.

response functions also differs between subsamples. For instance, with the LPJ-GUESS model, temperature increases have a negative mildly convex impact on yields in the Oxisols subsample, which are mostly located in the Tropics and subject to high summer temperatures and precipitations. Conversely, for most other soil order subsamples, the effect of temperature is concave. One should note that the confidence intervals for a few cases are relatively large (e.g. *Tmean_2* of rice with LPJ-GUESS in the Vertisols subsample) and therefore projection accuracy using the emulator for these soil subsample-GGCM combinations are likely to be less accurate than those for other subsamples.

The effect of CO_2 on crop yields, both direct—as captured by the non-linear representation of CO_2 and the interaction term between CO_2 and temperature—and indirect (via water use efficiency improvements)—as captured by the interaction term between CO_2 and precipitation—are accounted for in the regression and presented in Appendix E. The estimates indicate a concave relationship between CO_2 and yields for most GGCMs.

To evaluate the accuracy of the emulators, for each crop, GGCM and specification, the root mean square error (RMSE) is calculated to estimate the average error between predicted and 'actual' yields. For the S1fpintsoil specification, estimated for each soil subsample, the overall average RMSEs are weighted by the number of observations in each subsample. To account for differences in yield levels between crops and models, the normalized root mean square error (NRMSE) is calculated by dividing the RMSE by the difference between maximum and minimum yields. These goodness of fit measures are represented in Fig. 7. The bar graphs show that across crop models, the largest errors in absolute terms are associated with the PEGASUS and pDSSAT models. However, in relative terms, errors are usually the smallest for the PEGASUS and pDSSAT models and the largest for the GEPIC model. Modelling errors are generally larger for maize than for the three other crops, except for soybean for the GEPIC model. When considering functional forms, the average errors between the yields from the statistical models and the GGCMs are generally the lowest for the S1fpint specification. On average, the fractional polynomial functional form reduces errors by 0.04t/ha for maize, 0.03t/ha for rice and soybean, and 0.02t/ha for wheat compared to a quadratic function. The overall gain in goodness-of-fit compared to a higher-degree polynomial function is less clear-cut with RMSEs for the S1fpint specification on average 0.005t/ha less than for the S1polyint specification. Nonetheless, over all crops and crop models, the small gains in goodness of fit combined with the improved representation of the precipitation and temperature effects on crop yields, demonstrate that the fractional polynomial transformation is best suited to the task. When re-estimating the S1fpint specification at the soil order level (S1fpintsoil), the goodness of fit is by allowing parameter heterogeneity across soil order subsamples. Specifically, in the S1fpintsoil specification at the global level, errors are reduced by up to 0.09t/ha for maize, 0.02t/ha for rice, 0.01t/ha for soybean, and 0.09t/ha for wheat relative to the S1fpint specification. Overall, the gain in accuracy of the estimation involving soil types considerations appears substantial enough to warrant the additional level of complexity.

Table 6
Pearson's correlation between changes in crop yields from 2000 s to 2030s, 2050s, 2070 s and 2090 s between GGCMs and statistical emulators (S1fpintsoil specification).

Crop	Model	Pearson's Correlation Coefficient			
		2030s	2050s	2070s	2090s
Maize	GEPIC	0.48*	0.53*	0.48*	0.56*
	LPJ-GUESS	0.63*	0.69*	0.72*	0.73*
	LPJmL	0.66*	0.72*	0.76*	0.79*
	PEGASUS	0.47*	0.56*	0.63*	0.62*
	pDSSAT	0.43*	0.57*	0.64*	0.70*
Rice	GEPIC	0.51*	0.56*	0.51*	0.59*
	LPJ-GUESS	0.54*	0.60*	0.62*	0.64*
	LPJmL	0.53*	0.56*	0.56*	0.59*
Soybean	GEPIC	0.44*	0.50*	0.51*	0.60*
	LPJ-GUESS	0.63*	0.68*	0.70*	0.72*
	LPJmL	0.59*	0.65*	0.68*	0.70*
	PEGASUS	0.59*	0.65*	0.68*	0.70*
Wheat	GEPIC	0.56*	0.58*	0.55*	0.50*
	LPJ-GUESS	0.58*	0.59*	0.63*	0.63*
	LPJmL	0.43*	0.47*	0.48*	0.47*
	PEGASUS	0.43*	0.47*	0.48*	0.47*
	pDSSAT	0.36*	0.48*	0.58*	0.60*

Note: * $p < 0.1$.

4. Validation

To assess the ability of the statistical models to predict crop yields simulated by GGCMs, emulated yields are first compared with GCCM yields in an in-sample forecasting exercise considering the full sample. An out-of-sample validation exercise is then conducted by comparing emulated yields based on a partial sample excluding simulation from one climate model, and comparing emulated yields to GGCM yields from the excluded sub-sample. The validation analyses focus on the preferred specification, S1fpintsoil. Results for the S1fpint specification estimated at the global level are presented in Appendix G.

4.1. In-sample validation

To validate the emulators' prediction accuracy, the within-sample validation exercise is based on yield estimates using the full sample and predictions for each grid cell, year and climate model. Annual yields for each crop, GGCM and statistical emulator averaged over the three climate models and all grid cells for the whole globe are reported in the left panels of Fig. 8. The panels on the right represent the same global average but weighted by crop-specific irrigated area from the MIRCA2000 dataset (Portmann et al., 2010). The dark lines represent simulations from the emulator using the S1fpintsoil specification and the lighter lines are representative of the GGCMs' projections. The graphs show that average yields levels differ across GGCMs, despite being driven by the same climate data.[13] On average, however, predictions from the statistical emulators follow the same trend as projections from GGCMs. Some inter-annual yield variability is also captured by the statistical models, although with less accuracy. When considering yield projections weighted by harvested area, the graphs show on average much larger yields and greater inter-annual discrepancies between the emulators and the GGCMs, especially for maize at the beginning and end of the sample.

As a large share of the absolute crop yields estimated using the emulator is captured by the grid cell fixed effects, we compare temporal changes in crop yields estimated by the GGCMs and the emulators as an additional validation exercise. To this end, the Pearson's coefficients of correlation between predicted yield changes from the 2000 s (period 2000–2009) and four future decades (2030s, 2050s, 2070 s and 2090s) are reported in

Table 6 The correlation coefficients are all statistically significant at the 0.01 level, and overall the smallest coefficients are observed for the 2030 s period, and the largest coefficients are observed for the 2090 s period. For instance, the coefficient of correlation between changes in crop yields from the 2000 s to the 2030 s for wheat with the pDSSAT model is 0.36, indicating that the emulator capture only 13% of the changes in yields estimated by this crop model. In contrast, the coefficient of correlation between changes in crop yields from 2000 s to 2090 s for maize with the LPJmL model is 0.79, which indicates that 62% of the variation in the CGCM projections is captured by the emulator. From these results, the emulators appear better suited for replicating long-term crop yield changes than short-term variations. Additionally, a further indication of the benefits of estimating separate weather response functions for each soil subsample is that coefficients of correlation between changes in crop yield changes at the global level in the S1fprint specification are higher than those in the S1fpint specification for all cases (see Appendix G, Table G1).

Corresponding density-based plots comparing the changes in crop yields projected by the S1fpintsoil specification and the GGCMs from the 2000 s and the 2090 s are provided in Fig. 9 (plots for all periods are provided in Figures F18 to F21 in Appendix F). The plots show that for most models, a large share of the data points are located within −1 and +1 t/Ha, except for the LPJ-GUESS models, for which changes in yields are more widespread and are mostly positive. Generally, data points are clustered along the 45°white line, indicating a general agreement between changes in yields predicted with the GGCMs and those projected with the emulators.

To identify spatial patterns of agreements between climate change impact projections from the two types of models, left panels of Figs. 10–13 present the percentage changes in crop yield from 2000 s to 2090 s for each GGCM and sS1fpintsoil specification at the grid cell level. To compare the differences in impacts between the two approaches, the logarithmic ratios of percentage changes are calculated and presented in the right panels.[14] To focus

[13] Note that the LPJ-GUESS model simulate potential yields (yield non-limited by nutrient or management constraints) whereas the other crop models simulate actual yields.

[14] The log ratio is deemed more appropriate than a simple ratio which is not a symmetric measure. For example, if the GGCM projects an increase twice as large as

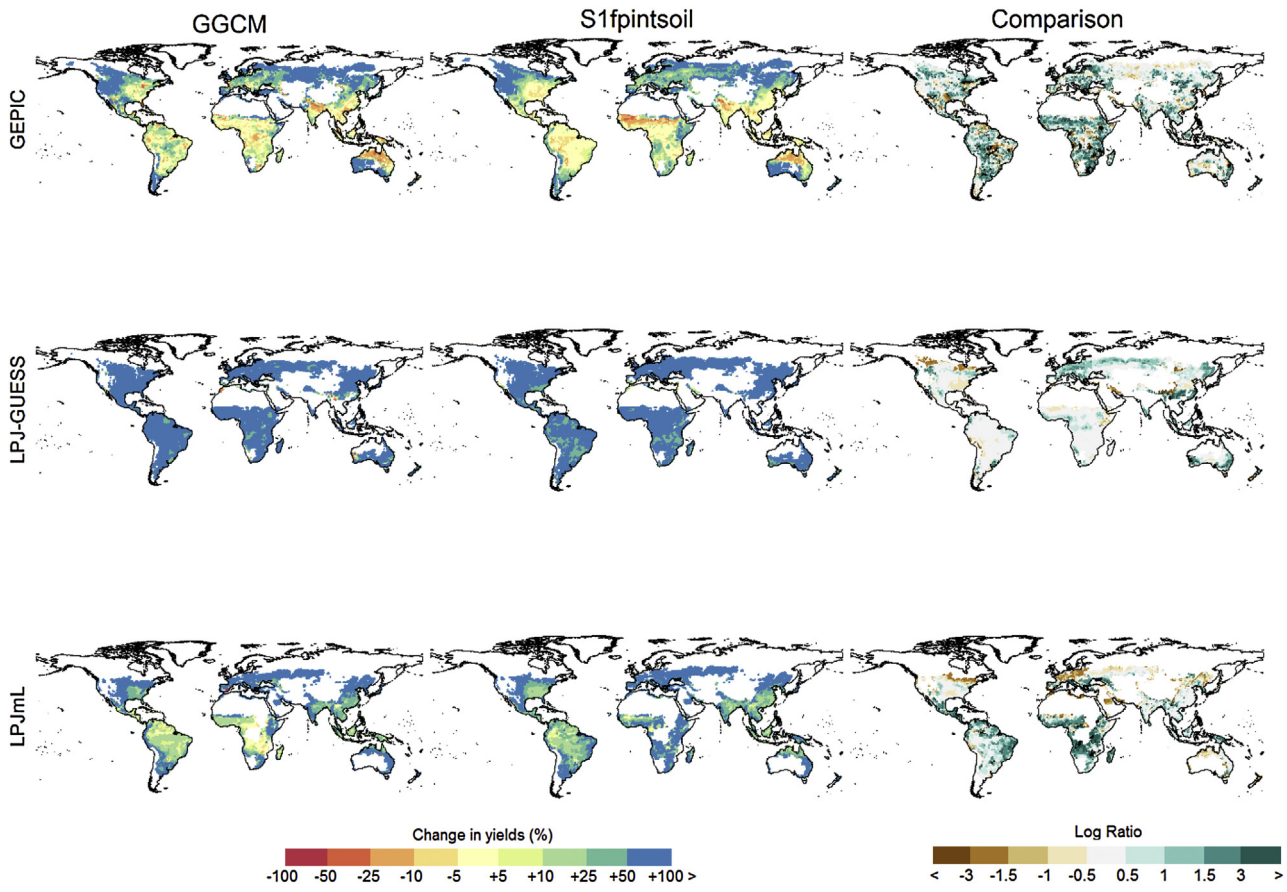

Fig. 11. Changes in rice yields from 2000 s to 2090 s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio).
Note: See note of Fig. 10.

on favorable crop growing regions, marginal growing areas (where yields projections from crop models are less than 1t/ha) are masked in white. Maps representing yield projections in levels form the S1fpintsoil specification and GGCMs and their differences (in absolute and percentage terms) are provided in Appendix F (Fig. F1 to F17).

In general, the maps confirm that the emulator reproduces the spatial patterns of climate change impacts on crop yields. Areas of disagreement regarding the sign of climate change impact on yields (represented in black on the comparison maps) are limited and generally observed in areas where the projected impact is close to zero. For all crops and GGCM, these areas of disagreement are also greatly reduced when considering the soil subsample emulator (S1fpintsoil) compared to the global emulator (S1fpint) (see Figure G20 to G23 in Appendix G).

For maize, the large increases in crop yields in the northernmost regions and decrease in the tropical regions by the end of the century projected by most GGCMs are reproduced by the emulator. However, as shown in the right panels, the emulator tends to underestimate the strength of the impacts projected by the LPJml, PEGASUS and pDSSAT models in northern America. For the LPJ-GUESS model, strong agreement between the GGCM and the emulator is observed over most of the Northern Hemisphere, but disagreement regarding the sign of the impacts is observed in central Africa where climate change impacts are relatively small. In all other GGCMs, however, yields changes projected with the emulator are overestimated in northern Eurasia.

For rice, spatial patterns of changes in crop yields projected by the GGCMs diverge significantly between GGCMs, with GEPIC projecting decreases in yields in Eastern US and the tropical regions, LPJ-GUESS projecting increase globally and LPJmL projecting small increases in tropical regions. Those impacts are also replicated by the emulators, although the emulator of the GEPIC model tends to overestimate yields in southeast US where rice productivity is projected to be high while the emulators for the two other models are more accurate. Large disagreements between the two methods are also observed in eastern South America for the LPJmL model.

For soybean, generally good spatial agreement of yield changes from the 2000 s to the 2090 s is observed between the emulator and GGCMs projections. In the southeast US, the emulator for the GEPIC model underestimates yield increases. For LPJmL, soybean yields changes are underestimated in northern America and underestimated in eastern Europe. For the LPJ-GUESS model, discrepancies are overall relatively small.

Wheat yield changes projected by the emulators are also in high agreement with those from the LPJ-GUESS model, except for small pockets of relatively high yields in South America, central Asia and southeast Australia. For the GEPIC and LPJmL models, however, the emulators tend to underestimate wheat yield changes in Europe and North America. For the pDSSAT model, the emulator overestimates yield changes in northern Eurasia where yields are relatively low. Additionally, the emulator projects detrimental impacts of climate change in eastern China where the PEGASUS model projects large beneficial impacts.

the emulator, say +20% vs +10%, the ratio would be 0.5. Now, if the GGCM projects an increase twice as small as the emulator, say +10% vs +20%, the ratio would be 2. The log ratio would be −0.3 and +0.3 respectively, indicating that the emulator underestimate climate change impacts in the first case and overestimates it in the second case in equal measures.

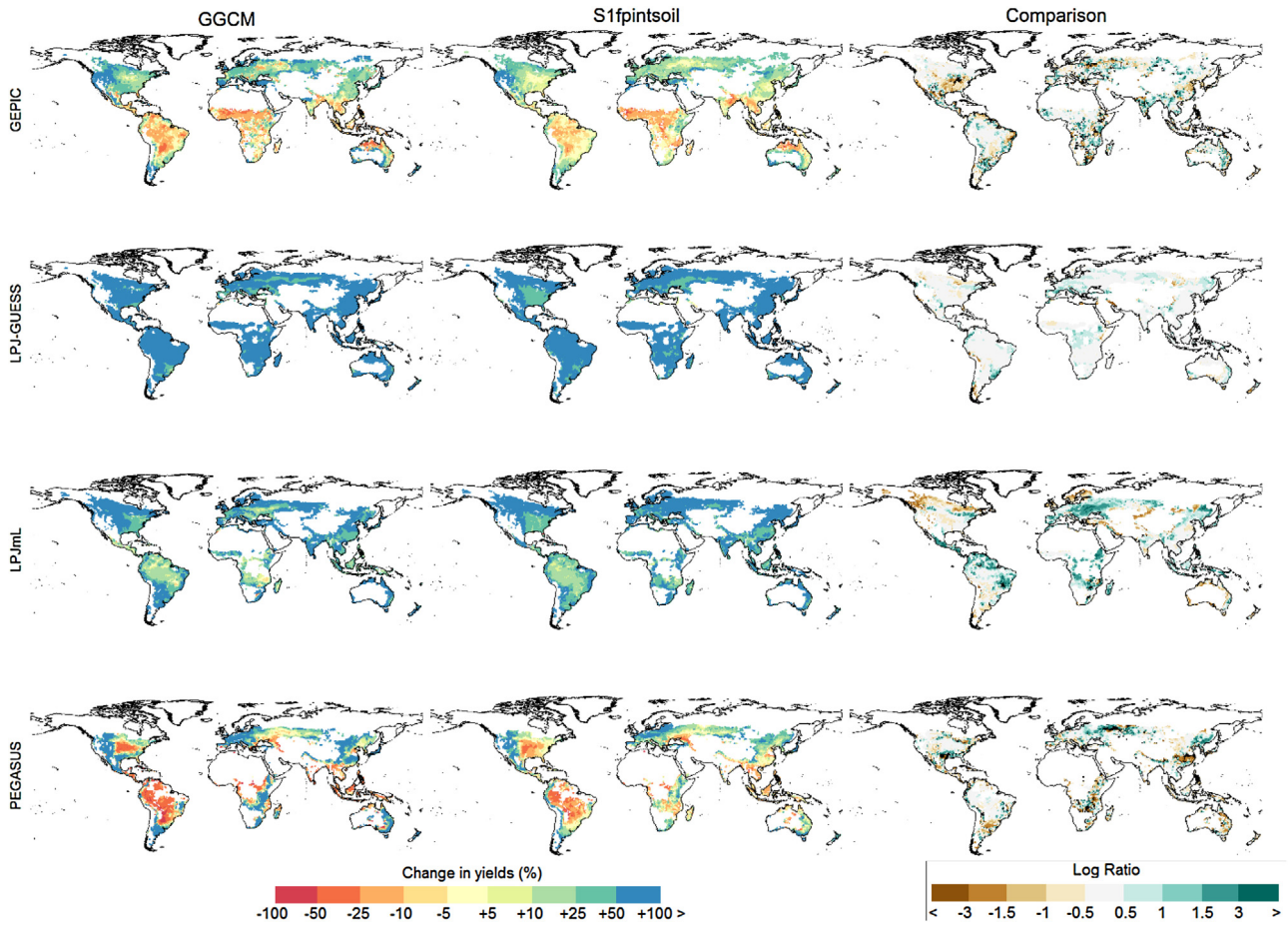

Fig. 12. Changes in soybean yields from 2000 s to 2090 s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio). *Note*: See note of Fig. 10.

4.2. Out-of-sample validation

To validate the ability of the emulator to perform under unknown climate change scenarios, an out-of-sample validation exercise is conducted by estimating the regression coefficients using a partial sample that includes data from all but one climate model, and then using these coefficients to emulate yields under weather variables estimated by the excluded climate model. This leave-one-GCM-out validation procedure is implemented three times in order to assess the predictive ability of the emulators for each omitted GCM.

Performance statistics for the out-of-sample validation exercise are reported in Table 7 and compared to the same statistics calculated using the full sample from Section 3. As expected, prediction errors are larger out-of-sample than in-sample but the differences are relatively small. The NRMSEs show a differential between the overall out-of-sample and the in-sample statistics ranging between 0.002 points for wheat with the PEGASUS model, and 0.008 points for maize with the LPJ-GUESS model.

To evaluate discrepancies between yield projections from GGCM yields and out-of-sample yields from the emulator over time, Fig. 14 shows the yield time series weighted by area harvested for each crop, GGCM and leave-one-GCM-out combination. The figures indicate that the emulated crop yields are generally underestimated for the NorESM1-M model when this GCM is excluded from the training dataset. This result is explained by the fact that yield projections under weather conditions from this model are, in most cases, higher than under other GCMs. Conversely, crop yield predictions from the statistical emulators tend to be over-estimated when the GFDL-ESM2 M model is excluded from the training sample, and for which yields are usually the smallest.

The conclusions from the out-of-sample validation exercise are in line with expectations from statistical analyses, which estimate an average effect. They further highlight the importance of considering the largest ensemble of climate change scenarios possible for the estimation of the emulators. In this regard, the wheat emulator for the PEGASUS model should be used with caution as it is trained on the HadGEM-ES GCM only. As the full sample was designed to encompass the extreme ranges of climate change currently being projected, statistical models estimated using this sample are therefore expected to provide the best predictions of crop yields even under alternative climate change scenarios—which are expected to be within the range of scenario considered.

5. Concluding remarks

This analysis provides simple emulation tools facilitating the assessment of climate change impacts on crop yields. The emulators are constructed based on an ensemble of crop yield simulations from five GGCMs as part of the ISI-MIP Fast Track intercomparison exercise. These GGCMs estimate the impact of weather on crop yields at a $0.5 \times 0.5^{\circ}$ resolution under various climate change scenarios. Based on a panel of crop yield and weather data at the grid-cell level, crop-specific response functions are estimated for each GGCM.

Building on Blanc and Sultan (2015), this analysis provides estimates for four crops: maize, rice, soybean and wheat. It focuses on a regression specification that include temperature and pre-

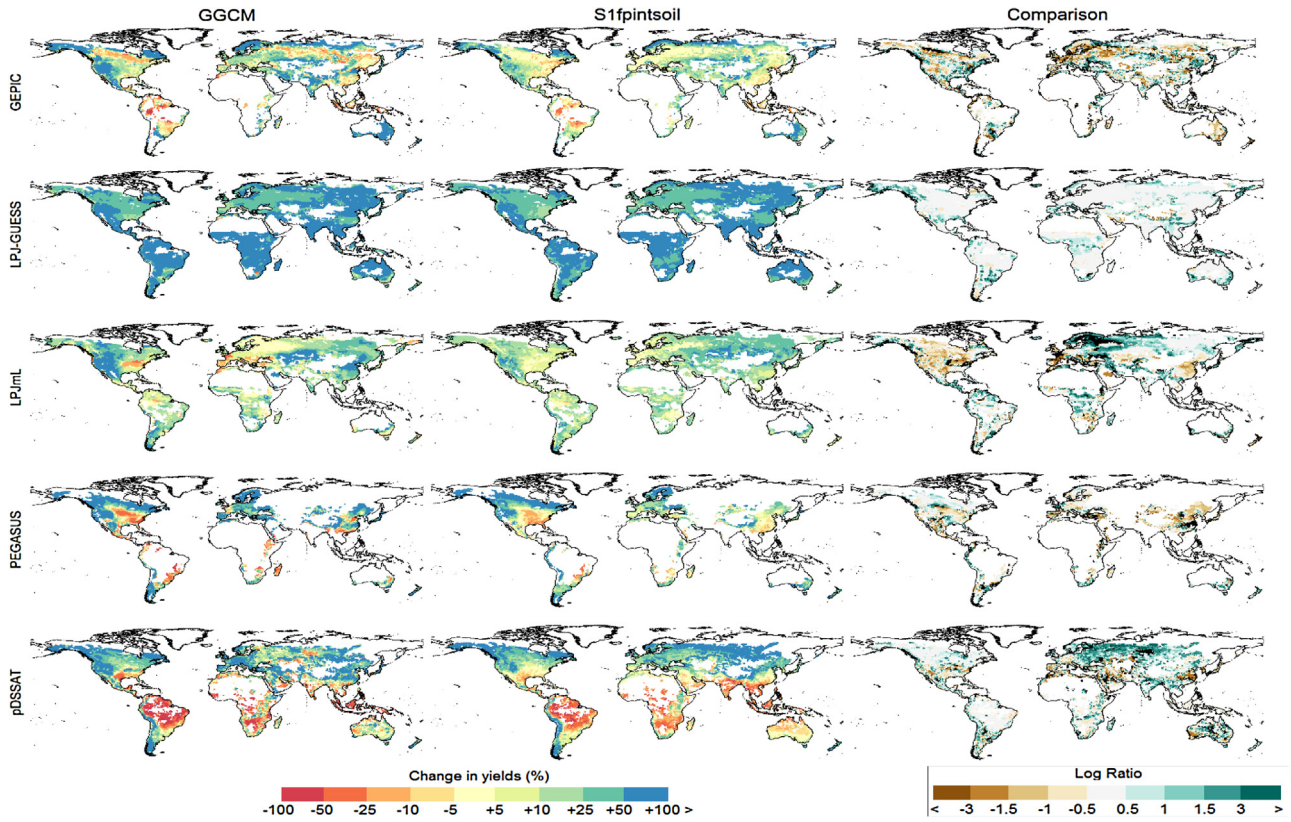

Fig. 13. Changes in wheat yields from 2000 s to 2090 s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio). *Note*: See note of Fig. 10.

cipitation only, which is deemed the best compromise in term of predictive ability and simplicity. As an extension to Blanc and Sultan (2015), this analysis compares the traditional non-linear representations of weather effects on crop yields, the quadratic form, and higher degree polynomial with fractional polynomial transformations of weather variables. Fractional polynomial transformations relax the symmetrical relationship constraint imposed by the quadratic transformation while allowing non-parametric flexibility, and addresses the tail-end behavioral issue posed by higher degree polynomial transformations. Additionally, the effect of soil type is addressed by estimating the fractional polynomial specification separately for each soil order, which allows parameter heterogeneity across subsamples. These subsample estimates, despite adding some complexity to the estimation procedure, generally improve the accuracy of the statistical emulators.

The validation exercises show that crop yield predictions from the emulator are reasonably representative of those from most GGCMs. Predictions from the statistical emulators follow the same trend as projections from GGCMs, although inter-annual yield variability is captured with less accuracy. When considering changes in yields overtime, correlation analyses show that the emulators are better able to capture long-term changes in yields projected by GGCMs than short-term variations. Spatial analyses of these changes reveal that, overall, the emulators tend to capture the spatial patterns of climate change impacts on crop yields. However, disagreements appear regarding the strength of the impacts. The emulators tend to over- or under-estimate climate change impact on yields in different regions depending on the GGCM considered. Therefore, when using the emulators for regional impact assessments, care should be taken when choosing the ensemble of GGCM that best capture the impact projected by the underlying model.

Out-of-sample validations show that, as expected, prediction accuracy is reduced when the training sample excludes yield responses to weather variables outside the range of values used to estimate the model. It is therefore critical to estimate the statistical emulator using the largest sample available, which is designed to encompass the largest range of plausible changes in temperature and precipitation over the twenty-first century.

The crop yield emulators estimated in this study provide an accessible and reliable tool to estimate changes in crop yields under alternative plausible user-defined changes in climate. However, due to GGCM specificities, simulations are more suited to assess long-term trends in yields rather than inter-annual yield variability. The use of the emulator to estimate climate change impact on crop yields should follow the same principles. Also, as shown by the ISI-MIP simulations, the different GGCMs considered in this analysis do not necessarily agree on the extent of the impact of climate change on crop yields even under a similar scenario of climate change. As none of the GGCMSs is deemed better than another at projecting future crop yields, it is important to consider predictions from many models to account for crop yield modeling uncertainty. By providing yield emulators for several crop models, this study provides a computationally efficient method for researchers to consider modeling uncertainty in climate change impact assessments.

Acknowledgments

We thank Niven Winchester for helpful comments and suggestions. We acknowledge the modeling groups (listed in Appendix A, Table A1 of this paper) and the ISI-MIP coordination team for their roles in producing, coordinating, and making available the ISI-MIP model output. We gratefully acknowledge the financial support for this work from the U.S. Department of Energy, Office

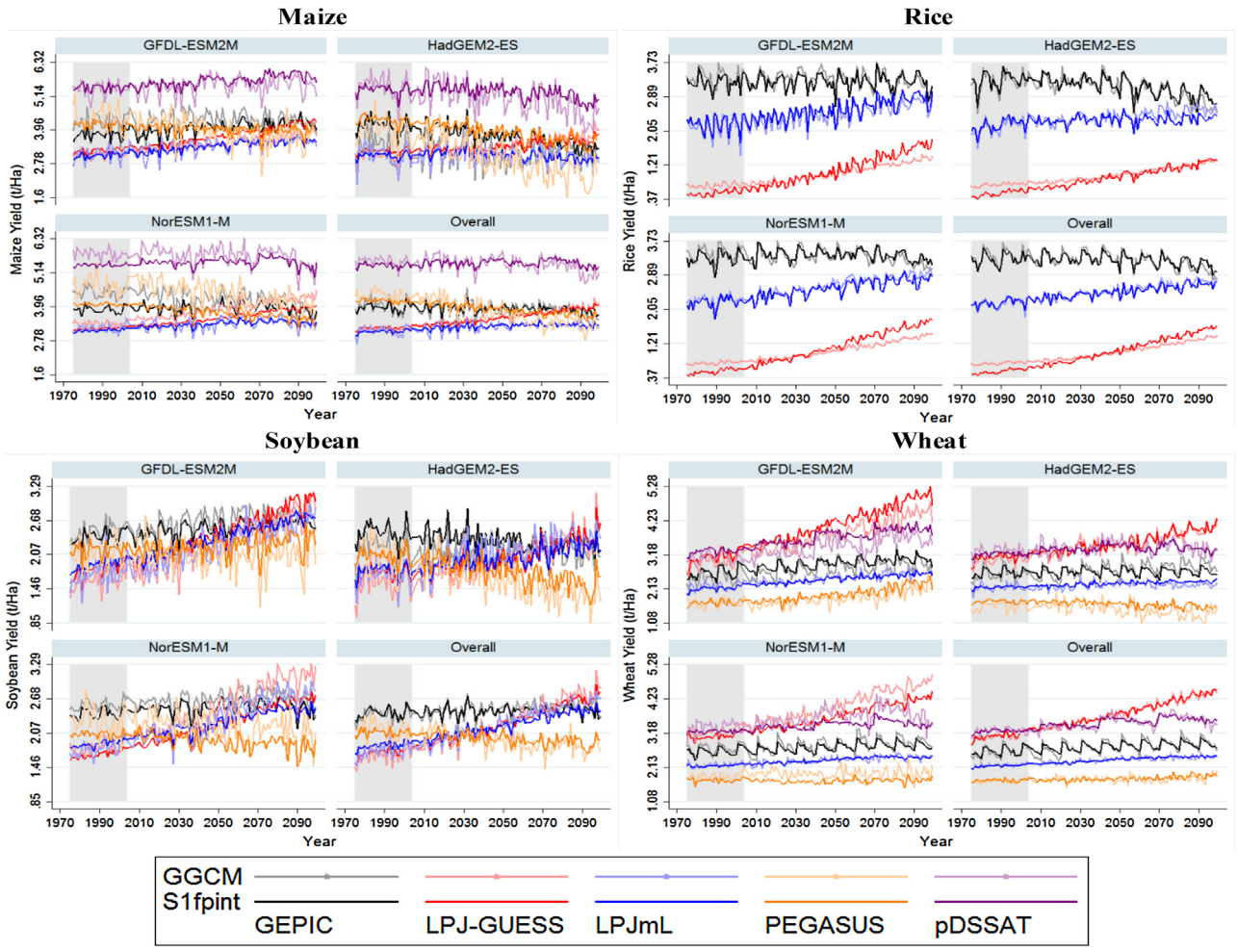

Fig. 14. Average crop yield projections weighted by area harvested from GGCMs and statistical models estimated using the S1fpintsoil specification in the leave-one-GCM-out validation exercise.

Table 7
RMSE and NRMSE statistics for the leave-one-GCM-out validation using the S1fpintsoil specification compared to the full sample.

Crop	Model	RMSE					NRMSE				
		GFDL−ESM2M	HadGEM2-ES	NorESM1-M	Overall	Full sample	GFDL−ESM2M	HadGEM2-ES	NorESM1-M	Overall	Full sample
Maize	GEPIC	0.96	0.99	0.83	0.93	0.84	0.07	0.07	0.06	0.06	0.06
	LPJ-GUESS	0.57	0.51	0.52	0.54	0.46	0.05	0.05	0.05	0.05	0.04
	LPJmL	0.77	0.77	0.70	0.75	0.69	0.04	0.04	0.04	0.04	0.04
	pDSSAT	1.43	1.34	1.34	1.37	1.27	0.06	0.06	0.06	0.06	0.05
	PEGASUS	1.32	1.37	1.34	1.34	1.25	0.04	0.04	0.04	0.04	0.04
Rice	GEPIC	0.79	0.73	0.66	0.72	0.66	0.06	0.05	0.05	0.05	0.05
	LPJ-GUESS	0.63	0.55	0.58	0.59	0.52	0.03	0.03	0.03	0.03	0.03
	LPJmL	0.62	0.63	0.54	0.59	0.55	0.03	0.03	0.03	0.03	0.03
Soybean	GEPIC	0.39	0.39	0.34	0.37	0.34	0.07	0.07	0.06	0.06	0.06
	LPJ-GUESS	0.41	0.35	0.36	0.37	0.34	0.04	0.03	0.03	0.04	0.03
	LPJmL	0.50	0.48	0.46	0.48	0.45	0.03	0.03	0.03	0.03	0.03
	PEGASUS	0.73	0.65	0.70	0.69	0.64	0.04	0.04	0.04	0.04	0.04
Wheat	GEPIC	0.60	0.58	0.53	0.57	0.53	0.05	0.05	0.05	0.05	0.05
	LPJ-GUESS	0.68	0.60	0.62	0.63	0.55	0.04	0.04	0.04	0.04	0.03
	LPJmL	0.48	0.45	0.42	0.45	0.39	0.03	0.03	0.02	0.03	0.02
	PEGASUS	0.74	0.73	0.70	0.72	0.68	0.03	0.03	0.03	0.03	0.02

Note: statistics for pDSSAT and wheat are not available as only simulations for the HadGEM2-ES scenario are available.

of Science under DE-FG02-94ER61937, the U.S. Environmental Protection Agency under XA-83600001-1 and other government, industry, and foundation sponsors of the Joint Program on the Science and Policy of Global Change. For a complete list of sponsors, please visit http://globalchange.mit.edu/sponsors/all.

Appendix A. Supplementary data

Supplementary materials associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.agrformet.2016.12.022.

References

Alexandrov, V.A., Hoogenboom, G., 2000. Vulnerability and adaptation assessments of agriculturalcrops under climate change in the Southeastern USA. Theor. Appl. Climatol. 67 (1–2), 45–63.

Bassu, S., et al., 2014. How do various maize crop models vary in their responses to climate change factors? Glob. Change Biol. 20 (7), 2301–2320.

Batjes, N.H., 2006. ISRIC-WISE Derived Soil Properties on a 5 by 5 Arc-minutes Global Grid, Wageningen, Netherlands, Available at http://www.isric.org.

Blanc, É., Strobl, E., 2013. The impact of climate change on cropland productivity: evidence from satellite based products at the river basin scale in Africa. Clim. Change 117 (4), 873–890.

Blanc, E., Sultan, B., 2015. Emulating maize yields from global gridded crop models using statistical estimates. Agric. For. Meteorol. 214–215, 134–147.

Blanc, É., 2012. The impact of climate change on crop yields in Sub-Saharan Africa. Am. J. Clim. Change 1 (1), 1–13.

Bondeau, A., et al., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Glob. Change Biol. 13 (3), 679–706.

Butt, T.A., McCarl, B.A., Angerer, J.A., Dyke, P.A., Stuth, J.W., 2005. The economic and food security implications of climate change in Mali. Clim. Change 68 (3), 355–378.

Challinor, A.J., et al., 2014. A meta-analysis of crop yield under climate change and adaptation. Nat. Clim. Change 4 (4), 287–291.

Deryng, D., Sacks, W.J., Barford, C.C., Ramankutty, N., 2011. Simulating the effects of climate and agricultural management practices on global crop yield. Glob. Biogeochem. Cycles 25 (2), GB2006.

Deryng, D., Conway, D., Ramankutty, N., Price, J., Warren, R., 2014. Global crop yield response to extreme heat stress under multiple climate change futures. Environ. Res. Lett. 9 (3), 034011.

Elliott, J., et al., 2013. The Parallel System for Integrating Impact Models and Sectors (pSIMS), Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery (XSEDE'13). Association for Computing Machinery, pp. 1–8.

FAO-UNESCO, 2005. Soil Map of the World, Digitized by ESRI. USDA-NRCS, Soil Science Division, World Soil Resources. FAO-UNESCO, Washington D.C.

Grassini, P., Eskridge, K.M., Cassman, K.G., 2013. Distinguishing between yield advances and yield plateaus in historical crop production trends. Nat. Commun. 4.

Haim, D., Shechter, M., Berliner, P., 2007. Assessing the impact of climate change on representative field crops in Israeli agriculture: a case study of wheat and cotton. Clim. Change.

Hempel, S., Frieler, K., Warszawski, L., Schewe, J., Piontek, F., 2013. A trend-preserving bias correction – the ISI-MIP approach. Earth Syst. Dyn. 4 (2), 219–236.

Holzkämper, A., Calanca, P., Fuhrer, J., 2012. Statistical crop models: predicting the effects of temperature and precipitation changes. Clim. Res. 51 (1), 11–21.

Jones, J., et al., 2003. The DSSAT cropping system model. Eur. J. Agron. 18 (3-4), 235–265.

Lindeskog, M., et al., 2013. Implications of accounting for land use in simulations of ecosystem services and carbon cycling in Africa. Earth Syst. Dyn. Discuss. 4, 235–278.

Liu, J., Williams, J.R., Zehnder, A.J.B., Yang, H., 2007. GEPIC – modelling wheat yield and crop water productivity with high resolution on a global scale. Agric. Syst. 94 (2), 478–493.

Lobell, D.B., Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agric. For. Meteorol. 150 (11), 1443–1452.

Lobell, D., Field, C., 2007. Global scale climate-crop yield relationships and the impacts of recent warming. Environ. Res. Lett. 2 (1), 1–7.

Mearns, L.O., Mavromatis, T., Tsvetsinskaya, E., Hays, C., Easterling, W., 1999. Comparative responses of EPIC and CERES crop models to high and low spatial resolution climate change scenarios. J. Geophys. Res.: Atmos. 104 (D6), 6623–6646.

Oyebamiji, O.K., et al., 2015. Emulating global climate change impacts on crop yields. Stat. Modell., in press.

Parry, M., Fischer, G., Livermore, M., Rosenzweig, C., Iglesias, A., 1999. Climate change and world food security: a new assessment. Glob. Environ. Change 9, 551–567.

Portmann, F.T., Siebert, S., Döll, P., 2010. MIRCA2000 – global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. Glob. Biogeochem. Cycles, 24.

Riahi, K., Grübler, A., Nakicenovic, N., 2007. Scenarios of long-term socio-economic and environmental development under climate stabilization. Technol. Forecasting Soc. Change 74 (7), 887–935.

Rosenzweig, C., Parry, M.L., 1994. Potential impacts of climate change on world food supply. Nature 367, 133–138.

Rosenzweig, C., et al., 2013. The agricultural model intercomparison and improvement project (AgMIP): protocols and pilot studies. Agric. For. Meteorol. 170 (0), 166–182.

Rosenzweig, C., et al., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proc. Natl. Acad. Sci. 111 (9), 3268–3273.

Royston, P., Altman, D.G., 1994. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. J. R. Stat. Soc. Ser. C: Appl. Stat. 43, 429–467.

Royston, P., Sauerbrei, W., 2008. Multivariable Model-building: A Pragmatic Approach to Regression Analysis Based on Fractional Polynomials for Modelling Continuous Variables. Wiley, Chichester, UK.

Schlenker, W., Lobell, D.B., 2010. Robust negative impacts of climate change on African agriculture. Environ. Res. Lett. 5, 1–8.

Schlenker, W., Roberts, M.J., 2009. Nonlinear temperature effects indicate severe damages to U. S: crop yields under climate change. Proc. Natl. Acad. Sci. U. S. A. 106 (37), 15594–15598.

Smith, B., Prentice, I.C., Sykes, M.T., 2001. Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space. Glob. Ecol. Biogeogr. 10 (6), 621–637.

Soil Survey Staff, 1999. Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys. Natural Resources Conservation Service, U.S. Department of Agriculture Handbook 436.

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An Overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. 93, 485–498.

Waha, K., van Bussel, L.G.J., Müller, C., Bondeau, A., 2012. Climate-driven simulation of global crop sowing dates. Glob. Ecol. Biogeogr. 21 (2), 247–259.

Warszawski, L., et al., 2014. The inter-Sectoral impact model intercomparison project (ISI-MIP): project framework. Proc. Natl. Acad. Sci. 111 (9), 3228–3232.

Williams, J., 1995. Chapter 25. the EPIC. In: Chapter 25. The EPIC, Computer Models of Watershed Hydrology. Water Resources Publications, Littleton, CO, pp. 909–1000.

Joint Program Reprint Series - Recent Articles

For limited quantities, Joint Program publications are available free of charge. Contact the Joint Program office to order.
Complete list: http://globalchange.mit.edu/publications

2017-1 Statistical emulators of maize, rice, soybean and wheat yields from global gridded crop models. Blanc, É., *Agricultural and Forest Meteorology,* 236, 145–161 (2017)

2016-25 Reducing CO_2 from cars in the European Union. Paltsev, S., Y.-H.H. Chen, V. Karplus, P. Kishimoto, J. Reilly, A. Löschel, K. von Graevenitz and S. Koesler, *Transportation*, online first (doi:10.1007/s11116-016-9741-3) (2016)

2016-24 Radiative effects of interannually varying vs. interannually invariant aerosol emissions from fires. Grandey, B.S., H.-H. Lee and C. Wang, *Atmospheric Chemistry & Physics,* 16, 14495–14513 (2016)

2016-23 Splitting the South: China and India's Divergence in International Environmental Negotiations. Stokes, L.C., A. Giang and N.E. Selin, *Global Environmental Politics*, 16(4): 12–31 (2016)

2016-22 Teaching and Learning from Environmental Summits: COP 21 and Beyond. Selin, N.E., *Global Environmental Politics*, 16(3): 31–40 (2016)

2016-21 Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. Armour, K.C., J. Marshall, J.R. Scott, A. Donohoe and E.R. Newsom, *Nature Geoscience* 9: 549–554 (2016)

2016-20 Hydrofluorocarbon (HFC) Emissions in China: An Inventory for 2005–2013 and Projections to 2050. Fang, X., G.J.M. Velders, A.R. Ravishankara, M.J. Molina, J. Hu and R.G. Prinn, *Environmental Science & Technology*, 50(4): 2027–2034 (2016)

2016-19 The Future of Natural Gas in China: Effects of Pricing Reform and Climate Policy. Zhang, D. and S. Paltsev, *Climate Change Economics*, 7(4): 1650012 (2016)

2016-18 Assessing the Impact of Typhoons on Rice Production in the Philippines. Blanc, É. and E. Strobl, *Journal of Applied Meteorology and Climatology,* 55: 993–1007 (2016)

2016-17 Uncertainties in Atmospheric Mercury Modeling for Policy Evaluation. Kwon, S.Y. and N.E. Selin, *Current Pollution Reports,* 2(2): 103–114 (2016)

2016-16 Limited Trading of Emissions Permits as a Climate Cooperation Mechanism? US-China and EU-China Examples. Gavard, C., N. Winchester and S. Paltsev, *Energy Economics,* 58(2016): 95–104 (2016)

2016-15 Interprovincial migration and the stringency of energy policy in China. Luo, X., J. Caron, V.J. Karplus, D. Zhang and X. Zhang, *Energy Economics,* 58(August 2016): 164–173 (2016)

2016-14 Modelling the potential for wind energy integration on China's coal-heavy electricity grid. Davidson, M.R., D. Zhang, W. Xiong, X. Zhang and V.J. Karplus, *Nature Energy,* 1: 16086 (2016)

2016-13 Pathways to Mexico's climate change mitigation targets: A multi-model analysis. Veysey, J., C. Octaviano, K. Calvin, S. Herreras Martinez, A. Kitous, J. McFarland and B. van der Zwaan, *Energy Economics,* 56(May): 587–599 (2016)

2016-12 Uncertainty in future agro-climate projections in the United States and benefits of greenhouse gas mitigation. Monier, E., L. Xu and R. Snyder, *Environmental Research Letters*, 11(2016): 055001 (2016)

2016-11 Impact of Aviation on Climate: FAA's Aviation Climate Change Research Initiative (ACCRI) Phase II. Brasseur, G., M. Gupta, B. Anderson, S. Balasubramanian, S. Barrett, D. Duda, G. Fleming, P. Forster, J. Fuglestvedt, A. Gettelman, R. Halthore, S. Jacob, M. Jacobson, A. Khodayari, K. Liou, M. Lund, R. Miake-Lye, P. Minnis, S. Olsen, J. Penner, R. Prinn, U. Schumann, H. Selkirk, A. Sokolov, N. Unger, P. Wolfe, H. Wong, D. Wuebbles, B. Yi, P. Yang and C. Zhou, *Bull. Amer. Meteor. Soc.*, 97(4): 561–583 (2016)

2016-10 Energy caps: Alternative climate policy instruments for China? Karplus, V.J., S. Rausch and D. Zhang, *Energy Economics*, 56(May 2016): 422–431 (2016)

2016-9 Cross-country electricity trade, renewable energy and European transmission infrastructure policy. Abrell, J. and S. Rausch, *Journal of Environmental Economics and Management*, 79(Sep 2016): 87–113 (2016)

2016-8 Transient Climate Impacts for Scenarios of Aerosol Emissions from Asia: A Story of Coal versus Gas. Grandey, B.S., H. Cheng and C. Wang, *Journal of Climate*, 29(8): 2849–2867 (2016)

2016-7 Climate and Land: Tradeoffs and Opportunities. Reilly, J.M. and J.M. Melillo, *Geoinformatics & Geostatistics: An Overview*, 4(1): 1000135 (2016)

2016-6 Projections of Water Stress Based on an Ensemble of Socioeconomic Growth and Climate Change Scenarios: A Case Study in Asia. Fant, C., C.A. Schlosser, X. Gao, K. Strzepek and J. Reilly, *PLoS ONE,* 11(3): e0150633 (2016)

2016-5 Determinants of Crop Yield and Profit of Family Farms: Evidence from the Senegal River Valley. Blanc, É., A. Lépine and E. Strobl, *Experimental Agriculture*, 52(1): 110–136 (2016)

2016-4 Metamodeling of Droplet Activation for Global Climate Models. Rothenberg, D. and C. Wang, *Journal of the Atmospheric Sciences*, 73, 1255–1272 (2016)

2016-3 Climate extremes and ozone pollution: a growing threat to China's food security. Tian, H., W. Ren, B. Tao, G. Sun, A. Chappelka, X. Wang, S. Pan, J. Yang, J. Liu, B.S. Felzer, J.M. Melillo and J. Reilly, *Ecosystem Health and Sustainability*, 2(1): 2332–8878 (2016)

MIT Joint Program on the Science and Policy of Global Change

Massachusetts Institute of Technology
77 Massachusetts Ave., E19-411
Cambridge MA 02139-4307 (USA)

T (617) 253-7492 F (617) 253-9845
globalchange@mit.edu
http://globalchange.mit.edu