![MIT Joint Program on the Science and Policy of Global Change logo]

Statistical Emulators of Irrigated Crop Yields and Irrigation Water Requirements

Élodie Blanc

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

Statistical Emulators of Irrigated Crop Yields and Irrigation Water Requirements

Élodie Blanc[1]

Abstract: This study provides statistical emulators of global gridded crop model included in the Inter-Sectoral Impact Model Intercomparison Project Fast Track project to estimate irrigated crop yields and associated irrigation water withdrawals simulated at the grid cell level. The ensemble of crop model simulations is used to build a panel of monthly summer weather variables and corresponding annual yields and irrigation water withdrawals from five gridded crop models. This dataset is then used to estimate crop-specific response functions for each crop model. In- and out-of-sample validation exercises confirm that the statistical emulators are able to replicate the crop models' spatial patterns of irrigated yields crop and irrigation water withdrawals reasonably well, both in levels and in terms of changes over time, although accuracy varies by model and by region. This study therefore provides a reliable and computationally efficient alternative to global gridded crop yield models.

1   Joint Program on the Science and Policy of Global Change. Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA. Email: eblanc@mit.edu

1. Introduction

The impact of climate change on crops can be assessed using process-based crop models (Boote *et al.*, 2013; Deryng *et al.*, 2014; Parry *et al.*, 1999; Rosenzweig and Parry, 1994a, 1994b; White *et al.*, 2011), statistical models (Auffhammer and Schlenker, 2014; Blanc and Strobl, 2013; Haim *et al.*, 2007; Hsiang, 2016; Lobell and Field, 2007; Schlenker and Roberts, 2009) or a combination of both (Roberts *et al.*, 2017) (i.e. a process model with parameters statistically estimated using historical observations). These models can then be included in Integrated assessment models (IAMs) which, by considering socio-economic and natural sciences mechanisms, provide a better representation of the agricultural sector. Calvin and Fisher-Vanden (2017) find that combining statistical or process-based models within IAMs helps predict climate change impacts on crop yields more accurately than on their own. Alternatively, the implementation of statistical emulators—statistical models trained on the outputs of a process-based model to capture the response functions from complex, computationally demanding and sometimes proprietary process-based crop models—in IAMs can help account for feedback loops from the agricultural sector (Ruane *et al.*, 2017) and can help account for modeling uncertainty (Monier *et al.*, 2018).

Statistical emulators have been used by Holzkämper, Calanca, and Fuhrer (2012) and Lobell and Burke (2010) to assess the capacity of statistical models to predict out-of-sample crop yields. Other studies have used emulated response functions to compare statistical and process based models for 'diagnostic purposes' (Lobell and Asseng, 2017; Schauberger *et al.*, 2017; Moore *et al.*, 2017). Crop yield emulators have also been developed to provide climate change impact assessment tools. Oyebamiji *et al.* (2015) provides crop yield emulators at the global level for five different crops but only considers one process-based crop model. Blanc and Sultan (2015) consider only maize but for five climate models. The scope of these emulators has then been expanded to three other crops (Blanc, 2017) and to both climate and crop models (Ostberg *et al.*, 2018). While Oyebamiji *et al.* (2015) and Ostberg *et al.* (2018)'s studies include emulators of irrigated crop yields, they don't consider the irrigation water requirements. However, as water availability may pose serious constraints to irrigation (Blanc *et al.*, 2017; Elliott *et al.*, 2014), water necessary to irrigate those crop is also a concern when estimating climate change impact on agriculture. This study proposes to fill this gap by developing statistical emulators of global gridded crop models for irrigated crops yields as well as the associated irrigation water withdrawals.

Building on Blanc and Sultan (2015) and Blanc (2017), the statistical emulators developed in this study are estimated based upon an ensemble of global gridded crop models (GGCM) simulations from the Inter-Sectoral Impact Model

Intercomparison Project (ISI-MIP) Fast Track experiment (Rosenzweig *et al.*, 2013; Warszawski *et al.*, 2014). This project was designed to compare GGCMs simulations, all driven by the same bias-corrected climate change projections obtained from the Coupled Model Intercomparison Project, phase 5 (CMIP5) simulations ensemble (Hempel *et al.*, 2013; Taylor *et al.*, 2012). In this study, the statistical emulators focus on irrigated crops and are estimated for maize, rice, soybean and wheat and five different GGCMs to provide an accessible tool for assessing the impact of climate change on irrigated crop yields and irrigation water withdrawals, while accounting for crop modeling uncertainty. In combination with the statistical emulators of rainfed crop yields developed in Blanc (2017), these emulators enhance integrated assessment modeling by facilitating the estimation of the impact of climate change on, separately, rainfed and irrigated crops.

The remainder of this paper presents the data and methods used to statistically estimate the emulators in Section 2 and the results are presented and discussed in Section 3. Validation of the emulators, both in- and out-of-sample are presented in Section 4. Section 5 concludes.

2. Material and methods

2.1 Data

In this analysis, data are sourced from the ISI-MIP Fast Track experiment, an inter-model comparison exercise where different GGCMs were used to simulate annual crop yields and irrigation water withdrawals under the same set of weather and CO_2 concentration inputs taken from the CMIP5 climate simulations.[1] Using these data, a panel dataset of GGCMs output and atmospheric conditions is constructed for the period 1975–2099.

2.1.1 Weather and CO2

Weather data at a 0.5×0.5-degree resolution (about 50km2) used as input into each GGCM are obtained from the CMIP5 climate data simulations. A subset of climate simulations is selected to be representative of the broadest plausible range of future climate change. Three General Circulation Models (GCMs), HadGEM2-ES, NorESM1-M, and GFDL-ESM2M, are selected to be representative of respectively, high, medium and low levels of global warming (Warszawski *et al.*, 2014). Daily bias-corrected weather data generated by these GCMs are provided for the 'historical' period of 1975 to 2005 and the 'future' period of 2006 to 2099. For the 'future' period, only one greenhouse gas Representative Concentration Pathway is considered, the

---

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

RCP 8.5, which is consistent with the highest level of global warming compared to historical conditions.

Based on the daily precipitation, and daily minimum (*Tmin*) and maximum (*Tmax*) temperatures, monthly averages of precipitation (*Pr*) and mean temperature (*Tmean* = (*Tmin* + *Tmax*)/2) are calculated for each summer month. For ease of reference, weather variables for each summer month are denoted by numbers suffixes so that _1, _2, and _3 refer to, respectively, June, July and August in the Northern Hemisphere and December, January and February in the Southern Hemisphere. For each climate scenario considered, the corresponding CO_2 concentrations data are extracted from Riahi, Grübler, and Nakicenovic (2007)[2].

2.1.2 Irrigated crop yields

Simulated annual irrigated crop yields (*YIR*) in metric tons per hectare (t/ha) at a 0.5×0.5-degree resolution are obtained from the ISI-MIP Fast Track experiment for five GGCMs: (1) the Geographic Information System (GIS)-based Environmental Policy Integrated Climate (GEPIC) model (Liu *et al.*, 2007; Williams and Singh, 1995); (2) the Lund Potsdam-Jena managed Land (LPJmL) dynamic global vegetation and water balance model (Bondeau *et al.*, 2007; Waha *et al.*, 2012); (3) the Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS) with managed land model (Bondeau *et al.*, 2007; Linzdeskog *et al.*, 2013; Smith *et al.*, 2001); (4) the parallel Decision Support System for Agro-technology Transfer (pDSSAT) model (Elliott *et al.*, 2013; Jones *et al.*, 2003); and (5) the Predicting Ecosystem Goods And Services Using Scenarios (PEGASUS) model (Deryng *et al.*, 2011). Although these GGCMs differ in their representation of crop phenology, leaf area development, root expansion, nutrient assimilation, and yield formation, they all account for the effect of water, heat stress and CO_2 fertilization, and assume no technological change.[3] However, the LPJ-GUESS model simulate potential yields (yield non-limited by nutrient or management constraints) whereas the other crop models simulate actual yields. Other divergences and GGCM-specific periodic patterns of yield projections are discussed in Blanc and Sultan (2015).

Irrigation water withdrawals

Associated with irrigated crop yields projections, GGCMs report irrigation water demand, or potential irrigation water withdrawal (*PIRRWW*), in mm per year at a 0.5×0.5-degree resolution. As the crop models make different assumptions about the efficiency of irrigation, the reported *PIRRWW* is

harmonized across all models to obtain estimates of water directly available to the crop, i.e. no losses during conveyance and application. More specifically, the *PIRRWW* data provided by pDSSAT is multiplied by 0.75 for maize, soy and wheat. *PIRRWW* data provided by LPJmL are multiplied by grid specific project efficiencies applicable to all crops.[4] All other models assume an irrigation use efficiency of 100%.

Soil orders

To account for soil conditions, soil orders are extracted from the FAO-UNESCO (2005) Soil Map of the World at the 0.5×0.5-degree resolution. It uses the USDA soil taxonomy (Soil Survey Staff, 1999)[5] classifying soils on the basis of physical and chemical properties observed in situ (e.g. soil horizons, structure, texture, color) and inferred from environmental conditions (e.g., soil temperature and moisture regimes). Soils are grouped into 12 main soil orders (Gelisols, Histosols, Spodosols, Andisols, Oxisols, Vertisols, Aridisols, Ultisols, Mollisols, Inceptisols, and Entisols) as described in Blanc (2017).

2.1.3 Summary statistics

Globally, the sample for each crop-GGCM combination is composed of, on average, 15 million observations covering about 44,000 grid cells (see **Table 1**).[6] Simulations from the

---

4   The spatial file containing project efficiencies is available for download at https://www.isimip.org/documents/213/irrigation_project_efficiencies.nc.

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

6   In the final sample, grid cells for which there are less than 10 output observations after data cleaning are omitted.

Table 1. GGCMs summary information

Crop	Model	Observations	Grid Cells
Maize	GEPIC	16,176,798	44,902
	LPJ-GUESS	15,958,849	43,824
	LPJmL	16,696,167	45,597
	PEGASUS	11,427,653	43,301
	pDSSAT	13,221,217	42,877
Rice	GEPIC	16,277,183	45,312
	LPJ-GUESS	15,252,499	43,789
	LPJmL	16,721,941	45,236
Soybean	GEPIC	16,197,571	45,211
	LPJ-GUESS	15,538,632	43,422
	LPJmL	16,650,813	45,558
	PEGASUS	8,314,743	39,642
Wheat	GEPIC	16,468,355	45,326
	LPJ-GUESS	14,960,416	41,820
	LPJmL	16,859,028	45,724
	PEGASUS	11,839,747	43,387
	pDSSAT	13,484,362	43,073

---

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

3   See Rosenzweig *et al.* (2014) for more details regarding each model's processes. 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.

PEGASUS and pDSSAT models for rice and pDSSAT model for soybean are not available. For wheat, simulations by the pDSSAT model are only available for the HadGEM2 GCM.

Summary statistics for irrigated crop yields and irrigation demand are detailed in **Table 2** by GGCM and GCM. The global average of irrigated crop yields differs amongst crops with yields ranging from 1.8t/Ha for soybean to 3.5t/ha for maize. Across GGCMs, the largest variation is observed for wheat, which ranges from 1.73 t/ha for the PEGASUS model to 4.4t/ha for the LPJ-GUESS model. Regarding irrigation, soybean requires the least water on average (92.5mm/year) and rice the most (114mm/year). Across GCMs, average irrigation water withdrawals are the largest under the NorESM1_M scenario and the lowest under the GFDL_ESM2M scenario. Irrigation requirements vary

Table 2. Summary statistics by GGCM and GCM

Crop	GGCM	GFDL_ESM2M			HadGEM2_ES			NorESM1_M		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
		Irrigated crop yields (t/Ha), *YIR*								
Maize	GEPIC	3.22	0.00	14.74	3.07	0.00	13.01	3.27	0.00	13.24
	LPJ-GUESS	3.71	0.00	15.27	3.95	0.00	12.30	3.91	0.00	12.91
	LPJmL	3.17	0.00	26.81	3.38	0.00	30.40	3.30	0.00	26.66
	PEGASUS	3.00	0.00	35.00	3.24	0.00	35.00	3.29	0.00	34.99
	pDSSAT	3.81	0.00	24.09	4.39	0.00	24.10	4.17	0.00	24.11
Rice	GEPIC	2.74	0.00	13.25	2.60	0.00	12.06	2.81	0.00	12.16
	LPJ-GUESS	2.13	0.00	20.69	2.19	0.00	22.84	2.29	0.00	20.83
	LPJmL	2.63	0.00	23.08	2.68	0.00	23.36	2.69	0.00	23.74
Soybean	GEPIC	1.38	0.00	5.89	1.33	0.00	6.06	1.41	0.00	6.30
	LPJ-GUESS	1.75	0.00	12.14	1.82	0.00	11.66	1.89	0.00	12.25
	LPJmL	1.99	0.00	19.47	2.06	0.00	19.66	2.08	0.00	20.73
	PEGASUS	1.98	0.00	22.21	2.17	0.00	23.74	2.21	0.00	22.52
Wheat	GEPIC	2.18	0.00	10.06	2.18	0.00	9.60	2.24	0.00	9.73
	LPJ-GUESS	4.36	0.00	24.12	4.35	0.00	22.69	4.53	0.00	22.28
	LPJmL	2.47	0.00	16.63	2.39	0.00	16.14	2.48	0.00	15.15
	PEGASUS	1.72	0.00	34.76	1.67	0.00	34.79	1.80	0.00	34.98
	pDSSAT	3.01	0.00	32.72	3.09	0.00	34.83	3.13	0.00	34.54
		Irrigation water withdrawals (mm), *PIRRWW*								
Maize	GEPIC	106.38	0.00	1068.00	109.37	0.00	935.10	90.65	0.00	914.30
	LPJ-GUESS	129.90	0.00	762.05	136.42	0.00	766.58	125.11	0.00	752.18
	LPJmL	156.16	0.00	1114.07	152.64	0.00	1119.02	150.24	0.00	1125.14
	PEGASUS	26.31	0.00	800.67	29.53	0.00	829.34	23.07	0.00	792.11
	pDSSAT	138.51	0.00	1000.50	150.69	0.00	1018.50	125.62	0.00	1055.25
Rice	GEPIC	143.09	0.00	1734.60	144.94	0.00	1630.20	122.79	0.00	1582.60
	LPJ-GUESS	85.43	0.00	1129.64	85.64	0.00	1059.34	81.13	0.00	1046.81
	LPJmL	150.54	0.00	962.42	149.43	0.00	993.82	145.71	0.00	970.27
Soybean	GEPIC	84.63	0.00	996.90	83.99	0.00	889.00	71.35	0.00	910.40
	LPJ-GUESS	121.51	0.00	968.59	126.60	0.00	1335.53	119.44	0.00	1306.70
	LPJmL	112.38	0.00	779.93	111.59	0.00	763.87	108.79	0.00	786.50
	PEGASUS	39.40	0.00	713.44	44.12	0.00	801.95	35.06	0.00	707.18
Wheat	GEPIC	117.64	0.00	721.80	107.13	0.00	639.50	103.61	0.00	612.10
	LPJ-GUESS	155.47	0.00	1059.80	159.79	0.00	1048.93	151.65	0.00	1052.28
	LPJmL	105.25	0.00	1047.35	99.09	0.00	979.18	103.79	0.00	951.06
	PEGASUS	20.87	0.00	712.25	23.93	0.00	722.45	18.89	0.00	718.73
	pDSSAT	138.40	0.00	2693.25	149.91	0.00	2797.50	141.12	0.00	2823.00

greatly across models, with the PEGASUS model simulating average irrigation water withdrawals below 40mm/year for all crops, whereas all other GGCMs (except GEPIC for soybean and LPJ-GUESS for rice) exceed 100mm/year.

Atmospheric CO_2 concentrations, which are the same for all GCM-GGCM combinations, range from 331 to 927 parts per million (ppm) between 1975 and the end of the century. Summary statistics of *Tmean* and *Pr*, and *CO_2* averaged over all GGCMs are presented in **Table 3**. On average, temperatures are the highest in the second month of summer and precipitation is the lowest in the first month of summer. Across GCMs, temperatures are the greatest under the HadGEM2-ES model and the lowest under the GFDL-ESM2M GCM, but no clear pattern of precipitation emerges amongst GCMs. Weather statistics details at the soil order level indicate that mid-summer temperature range between 18˚C in the Spodosols regions (acidic soils developing under coniferous vegetation) to 30˚C in the Vertisols regions (clay-rich soils in climates with distinct dry seasons). Precipitation ranges from less than 1mm/day in the Aridisols regions (prone to salinization and typical to arid regions) to more than 7mm/day in the Oxisols regions (mineral soils found in tropical and subtropical latitudes). More details regarding the weather variables statistics are available in Blanc (2017).

Table 3. Mean values of summer temperature and precipitation by GCM at the global level

Variable	Unit	GFDL_ESM2M	HadGEM2_ES	NorESM1_M
Tmean_1	˚C	21.4	22.8	22.0
Tmean_2	˚C	23.1	24.5	23.9
Tmean_3	˚C	22.4	23.8	22.9
Pr_1	mm/day	3.2	3.0	3.0
Pr_2	mm/day	3.5	3.5	3.5
Pr_3	mm/day	3.5	3.5	3.5

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

2.2 Methods

2.2.1 Specifications

Our methodology extends the work of Blanc and Sultan (2015) and Blanc (2017). In these studies, rainfed yields were estimated using a parsimonious specification that only included average summer precipitation and temperature weather variables, CO_2 concentrations, and interactions among these variables:

$$YRF_{lat,lon,gcm,y} = \alpha + \sum_{i=1}^{3} \beta_i Pr_{i_{lat,lon,gcm,y}} + \sum_{i=1}^{3} \theta_i Tmean_{i_{lat,lon,gcm,y}} + \vartheta CO2_{gcm,y} + \tag{1}$$

$$\sum_{i=1}^{3} \gamma_i Pr_{i_{lat,lon,gcm,y}} * Tmean_{i_{lat,lon,gcm,y}} + \sum_{i=1}^{3} \varepsilon_i Pr_{i_{lat,lon,gcm,y}} * CO2_{gcm,y} +$$

$$\sum_{i=1}^{3} \kappa_i Tmean\_i_{lat,lon,gcm,y} * CO2_{gcm,y} + \delta_{lat,lon} + \rho_{lat,lon,gcm,y}$$

where for each year, *y*, *YRF* corresponds to rainfed 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 mean precipitation and temperature variables for each summer month *i*. *CO_2* 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 to account for soil fertility erosion and CO_2 concentration for the pDSSAT and GEPIC models respectively.

For irrigated crops, the five GGCMs considered in this study assume that irrigation is applied to compensate for the lack of precipitation. More specifically, for the GEPIC model, *"full irrigation*

was set as a complete elimination of water stress of crops" (Rosenzweig *et al.*, 2014). In the four other models, however, irrigation is triggered when soil moisture is insufficient. More specifically, For the LPJ-GUESS and LPJmL models, *"additional water is provided as soon as the water content of the upper soil layer is insufficient"* (Bondeau *et al.*, 2007). The PEGASUS model ensures *"that soil is sufficiently moist to avoid water stress in irrigated land"* (Deryng *et al.*, 2011). The pDSSAT model, "*Determines daily irrigation, based on read-in values or automatic applications based on soil water depletion*" (Jones *et al.*, 2003). For these models, water stress may not necessarily be completely eliminated by full irrigation (Rosenzweig *et al.*, 2014).

Equation (1) is adopted to estimate yields for irrigated crops, *YIR*, using two specifications. The first specification, assumes that irrigation eliminates most water stress and therefore precipitation has no impact on crop growth and yields. The specification excluding precipitation can be specified as a function of temperature and CO_2, and corresponding interaction terms:

$$YIR_{lat,lon,gcm,y} = \alpha + \sum_{i=1}^{3} \theta_i Tmean_{i_{lat,lon,gcm,y}} + \vartheta CO2_{gcm,y} + \sum_{i=1}^{3} \gamma_i Tmean\_i_{lat,lon,gcm,y} *$$
$$CO2_{gcm,y} + \delta_{lat,lon} + \rho_{lat,lon,gcm,y} \tag{2}$$

To assess the effect of precipitation that may not have been completely eliminated by irrigation, a second specification including precipitation is specified as:

$$YIR_{lat,lon,gcm,y} = \alpha + \sum_{i=1}^{3} \beta_i Pr_{i_{lat,lon,gcm,y}} + \sum_{i=1}^{3} \theta_i Tmean_{i_{lat,lon,gcm,y}} + \vartheta CO2_{gcm,y} +$$
$$\sum_{i=1}^{3} \gamma_i Pr_{i_{lat,lon,gcm,y}} * Tmean_{i_{lat,lon,gcm,y}} + \sum_{i=1}^{3} \varepsilon_i Pr_{i_{lat,lon,gcm,y}} * CO2_{gcm,y} +$$
$$\sum_{i=1}^{3} \kappa_i Tmean\_i_{lat,lon,gcm,y} * CO2_{gcm,y} + \delta_{lat,lon} + \rho_{lat,lon,gcm,y} \tag{3}$$

Associated with each crop yield, GGCMS also provide annual irrigation water requirements (*PIRRWW*). Consistent with the methodology used to estimate crop yields, water demand for irrigation is estimated as a function of monthly weather and CO_2 concentrations. In a first specification, the weather vector is composed of mean monthly temperature and precipitation:

$$PIRRWW_{lat,lon,gcm,y} = \alpha + \sum_{i=1}^{3} \beta_i Pr_{i_{lat,lon,gcm,y}} + \sum_{i=1}^{3} \theta_i Tmean_{i_{lat,lon,gcm,y}} + \vartheta CO2_{gcm,y} +$$
$$\sum_{i=1}^{3} \gamma_i Pr_{i_{lat,lon,gcm,y}} * Tmean_{i_{lat,lon,gcm,y}} + \sum_{i=1}^{3} \gamma_i Pr_{i_{lat,lon,gcm,y}} * CO2_{gcm,y} +$$
$$\sum_{i=1}^{3} \gamma_i Tmean\_i_{lat,lon,gcm,y} * CO2_{gcm,y} + \delta_{lat,lon} + \rho_{lat,lon,gcm,y} \tag{4}$$

A second specification considers evapotranspiration (*ETo*) instead of temperature to account for the effect of summer weather on irrigation requirements:

$$PIRRWW_{lat,lon,gcm,y} = \alpha + \sum_{i=1}^{3} \beta_i Pr_{i_{lat,lon,gcm,y}} + \sum_{i=1}^{3} \theta_i ETo_{i_{lat,lon,gcm,y}} + \vartheta CO2_{gcm,y} +$$
$$\sum_{i=1}^{3} \gamma_i Pr_{i_{lat,lon,gcm,y}} * ETo_{i_{lat,lon,gcm,y}} + \sum_{i=1}^{3} \gamma_i Pr_{i_{lat,lon,gcm,y}} * CO2_{gcm,y} +$$
$$\sum_{i=1}^{3} \gamma_i Tmean\_i_{lat,lon,gcm,y} * CO2_{gcm,y} + \delta_{lat,lon} + \rho_{lat,lon,gcm,y} \tag{5}$$

As demonstrated by Blanc (2017), the weather effect on crops differs across soil types. The preferred estimation strategy therefore consists of estimating separate weather response functions for each soil order.[7]

2.2.2 Non-linear transformations

Weather variables are expected to have a non-linear effect on crop yields. Blanc (2017) find that a fractional polynomial specification is preferable to a quadratic an multinomial transformations frequently used in production functions as it relaxes the symmetry constraint imposed by quadratic terms but allows non-parametric flexibility from multinomial transformations. A fractional

---

[7]   In this analysis, response functions for the Gelisols soil order are not estimated, as this soil order represents soils permanently frozen.

polynomial model of degree m defining the relationship variables Y and X is defined as:

$$Y = \alpha_0 + \sum_{j=1}^{m} \alpha_j X^{(p_j)} + \mu \qquad (6)$$

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

$$X^{(p_j)} = \begin{cases} X^{p_j} & if \ p_j \neq 0 \\ lnX & if \ p_j = 0 \end{cases} \qquad (7)$$

where the term p_j is multiplied by another lnX for each repeated power. A closed-test algorithm performing a backward elimination (starting from the most complex specification) is used to fit the multivariable fractional polynomial model. For irrigated crop yields, following Blanc (2017), the maximum permitted degree is $m=2$. Regarding irrigation demand, the effect of temperature and precipitation is expected to be non-linear but without a turning point (i.e. it is expected that precipitation will reduce demand for irrigation water at a declining rate and will never lead to an increase in irrigation demand). To ensure this relationship, the maximum permitted degree of fractional polynomial transformation is limited to $m=1$.

Following Royston and Sauerbrei (2008), powers are chosen from among the set {-2, -1, -0.5, 0, 0.5, 1, 2, 3}. The specifications used to estimate crop yields and irrigation water requirements are summarized in **Table 4**.

3. Results

Based on the methodology presented Section 2, multiple specifications are estimated for both irrigated yields and irrigation demand. Results for irrigated crop yields and irrigation water requirements are presented in Section 3.1 and 3.2 respectively. The power terms used for the preferred specifications are reported in Appendix B and the regression results are presented in Appendix C. The corresponding estimated values for δ (the grid cell fixed effect) are provided in Appendix D.

3.1 Regression results for irrigated yields

For each crop and GGCM, regressions for irrigated yields are estimated for each specification S1 and S2 considering the fractional polynomial transformations at the soil order subsample level (S1fpintsoil and S2fpintsoil). As presented in **Figure 1**, the normalized root mean square error (NRMSE), which is calculated by dividing the RMSE by the difference between maximum and minimum yields, indi-

Table 4. Specification description

Dependent variable	Specification	Variables and non-linear transformations
YIR	S1fpintsoil	Tmean_p1, Tmean_p2, CO₂_p1, CO₂_p2
	S2fpintsoil	Pr_p1, Pr_p2, Tmean_p1, Tmean_p2, CO₂_p1, CO₂_p2
PIRRWW	S1fpintsoil	Pr, Tmean_p1, Tmean_p2, CO₂_p1, CO₂_p2
	S2fpintsoil	Pr, ETo _p1, ETo _p2, CO₂_p1, CO₂_p2

Note: suffix _sq denotes square terms, _p1 and _p2 power terms; All specifications include interaction terms and are estimated at the soil order level.
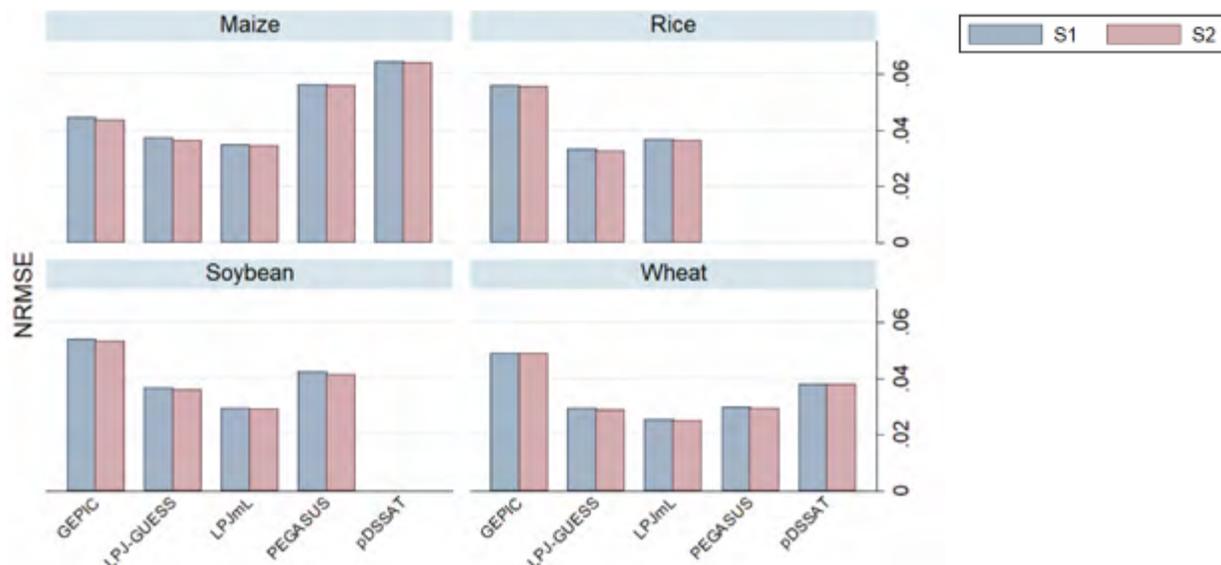

Figure 1. Goodness of fit of the irrigated yield statistical emulators by crop and GGCM (S1fpintsoil and S2fpintsoil specifications)

cates that the average error between predicted and 'actual' irrigated yields range from around 4% to 6% of maize and rice yields, 3% to 6% of soybean yields and 2% to 5% of wheat yields. Across GGCMs, the graph shows that lowest NRMSE are found for the LPJml and LPJ-GUESS models, while GEPIC has the highest NRMSE for all crops except maize. Examining differences across specifications, only slightly lower NRMSEs are found for the S2 specification across most crops and GGCMs. To favor simplicity, the most parsimonious S1 specification assuming that irrigation eliminates water stress (i.e. excluding the effect of precipitation) is thereafter preferred.

The S1fpintsoil regression results show that summer temperatures have a significant impact on irrigated yields from all GGCMs and crops. **Figure 2** provides an illustration of the average effect of temperature during the second month of summer while holding covariates at their mean values detailed for each soil sample. The figure shows that fractional polynomial transformation captures the non-linear effect of mid-summer temperature on irrigated crop yields, with in some cases, a negative skewness of the curve representing a sharp decrease in yields associated with high temperature. Similar to the results in Blanc (2017), the average effect of

temperature on crop yields differ depending on the soil order sample considered.

3.2 Regression results for irrigation water withdrawal

As for crop yields, regressions for irrigation water withdrawal are estimated for each crop and GGCM at the soil order subsample level considering both specifications S1 and S2 with fractional polynomial transformations (S1fpintsoil and S2fpintsoil). The NRMSE presented in **Figure 3**, shows an average error between predicted and 'actual' irrigation water withdrawals ranging between around 4% to 6% for most cases. Errors for *PIRRWW* for wheat with GEPIC reaches almost 8% whiles with pDSSAT they are closer to 2%. Across all crops and models, the NRMSE for the S1 specification is found to be slightly lower or equal to the S2 specification. The S1 specification is thereafter preferred.

As for irrigated crop yields, the regression results for the S1fpintsoil specification indicate that summer weather have a significant impact on irrigation water withdrawals from all GGCMs and crops. Illustrations of the average effect of temperature and precipitation during the second month of summer while holding covariates at their mean values
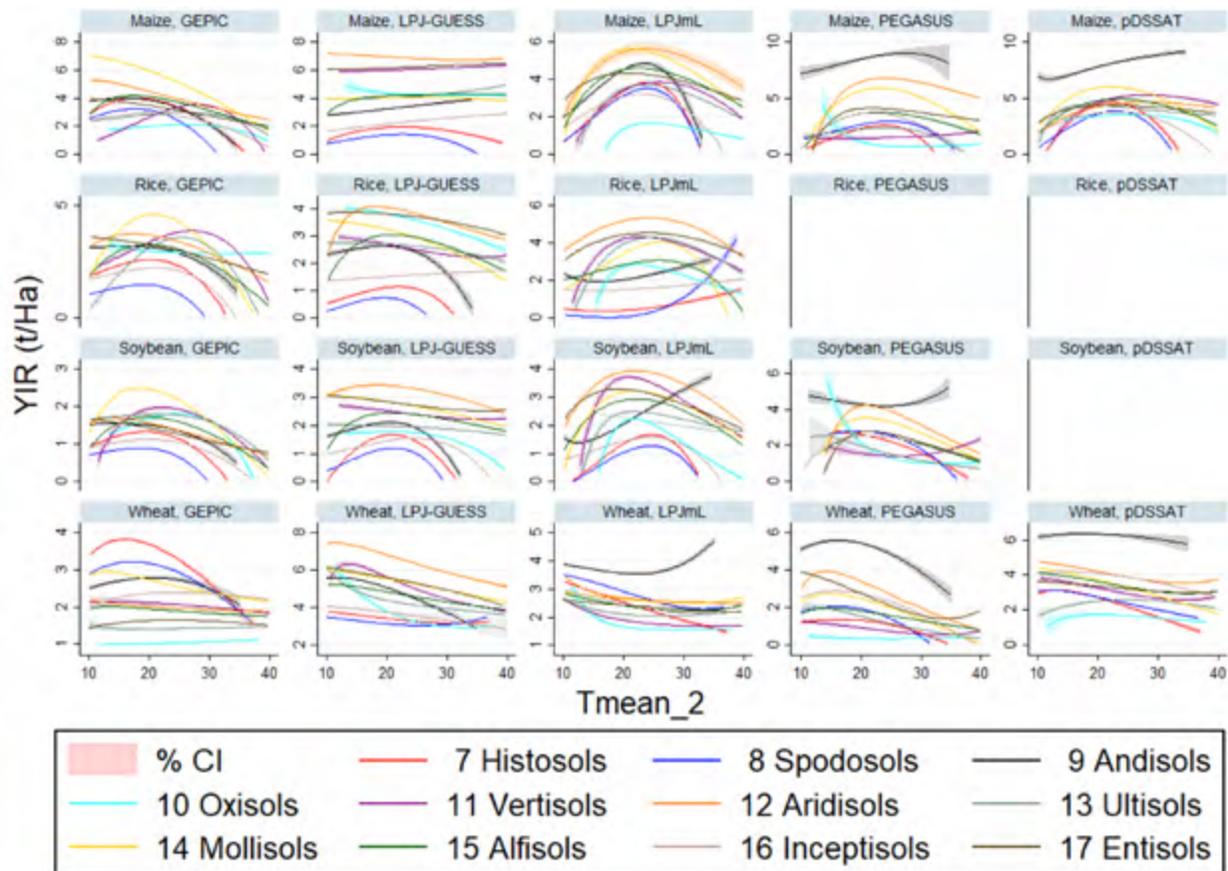
Figure 2. Effect of *Tmean_2* on *YIR* by crop and GGCM for the S1fpintsoil specification

Note: covariates are held at their mean values.
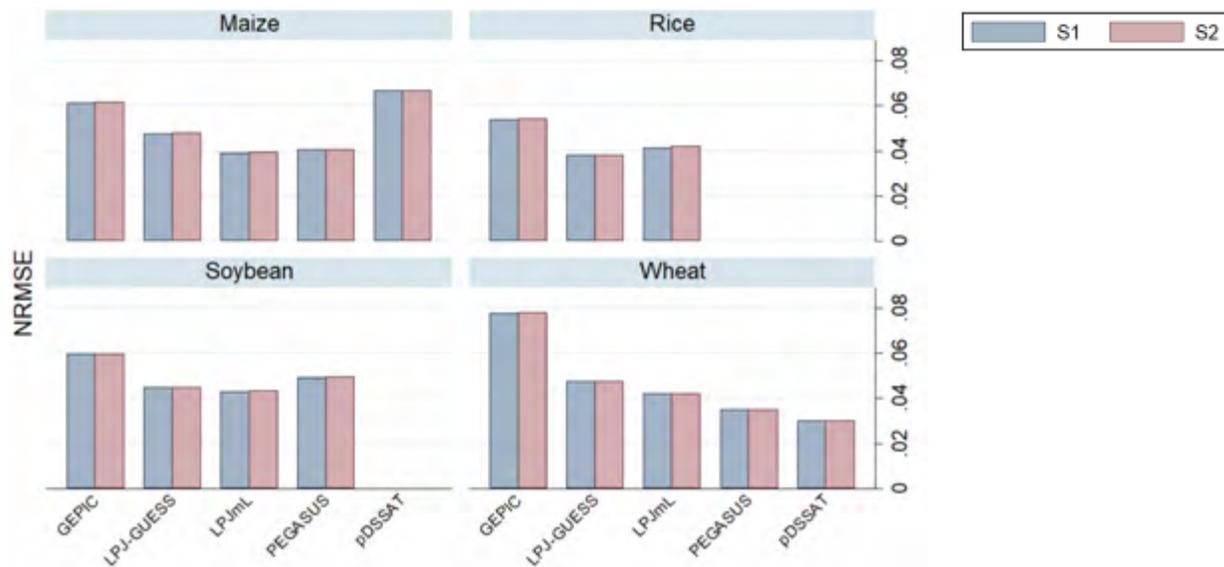
Figure 3. Goodness of fit of the irrigation water withdrawal statistical emulators by crop and GGCM (S1fpintsoil and S2fpintsoil specifications)
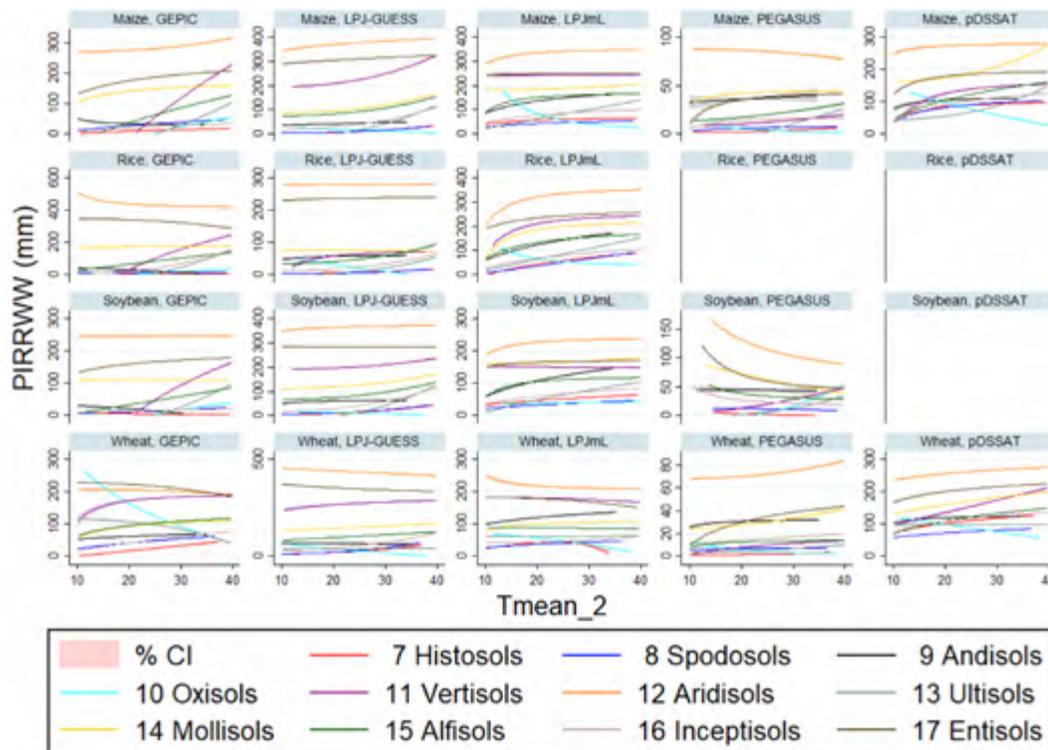
Figure 4. Effect of *Tmean_2* on *PIRRWW* by crop and GGCM for the S1fpintsoil specification

Note: covariates are held at their mean values.

detailed for each soil sample are provided in **Figure 4** and **Figure 5**. The figures indicate that the average effect of weather on irrigation water withdrawals varies by soil type. For instance, the effect of temperature and precipitation on irrigation water withdrawals is generally the largest in Aridisols regions, which are characteristic of arid regions. Figure 4 shows that temperature generally has a positive effect on irrigation water withdrawals, which is consistent with an increase in evaporation associated with higher temperature. Figure 5 indicates that in most cases, at low level of precipitation, an increase in rainfall is associated with a sharp decline in irrigation water withdrawals. The effect levels off when precipitation rates exceed around 2mm/day.
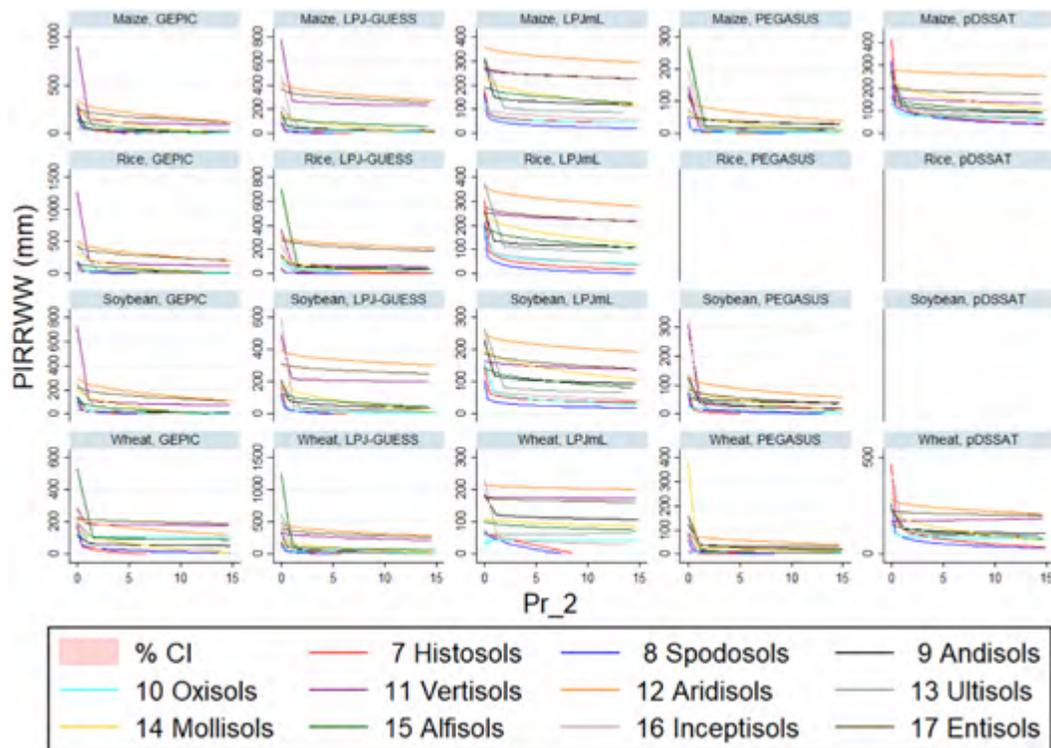
Figure 5. Effect of *Pr_2* on *PIRWW* by crop and GGCM for the S1fpintsoil specification

Note: covariates are held at their mean values.

4. Validation

To evaluate the accuracy of the statistical models at reproducing irrigated crop yields and associated irrigation water withdrawals simulated by GGCMs, the emulators' within- and out-of-sample projections are compared with those from GCCMs. Both validation exercises are lead using the preferred specification, S1fpintsoil.

4.1 In-sample validation

4.1.1 Irrigated crop yields

The within-sample validation exercise is performed on the full sample of irrigated yields estimates for each crop, grid cell, year, and climate model. To evaluate the emulators' prediction accuracy overtime, time series of average irrigated crop yields from GGCMs and statistical emulators are presented in **Figure 6**. The left hand side panels present annual irrigated yields for each crop averaged over the three climate models and all grid cells for the whole globe. Similar global averages but weighted by crop-specific irrigated harvested area (sourced from the MIRCA2000 dataset; Portmann *et al.*, 2010) are presented on the right hand side panels. The light colored lines represent the GGCMs' projections and the dark colored lines characterize simulations from the emulator (using the S1fpintsoil specification). The graphs show that, while average yields projections levels driven by the same cli-

mate data differ between GGCMs, predictions from the statistical emulators follow on average the same trend as projections from GGCMs, although inter-annual variability is captured with less accuracy. When focusing on irrigated areas, similar remarks can be made, except for irrigated yields of maize simulated with the pDSSAT model, rice with the LPJ-GUESS model, and soybean with the PEGASUS model where greater inter-annual divergences between the emulators and the GGCMs are observed at the beginning and at the end of the sample.

To assess the degree of spatial agreement between the emulator and the GGCMs, maps presenting climate change impact projections estimated by those models over the 2090s period are provided in Appendix E. The maps show that the emulators reproduce the spatial patterns of irrigated crop yields with reasonable accuracy. Similar spatial assessment maps considering the change in irrigated crop yields from 2000s to 2090s are provided in **Figure 7** to **Figure 10**. The first two columns of each figure represent at the grid cell level the percentage changes in irrigated crop yields for each GGCM and emulator (S1fpintsoil specification) respectively. The last columns of each figure show the logarithmic ratios of percentage changes to better distinguish the regions of divergence between the two models. Overall, the maps show that GGCMs project increases in irrigated crop yields poleward for most crops
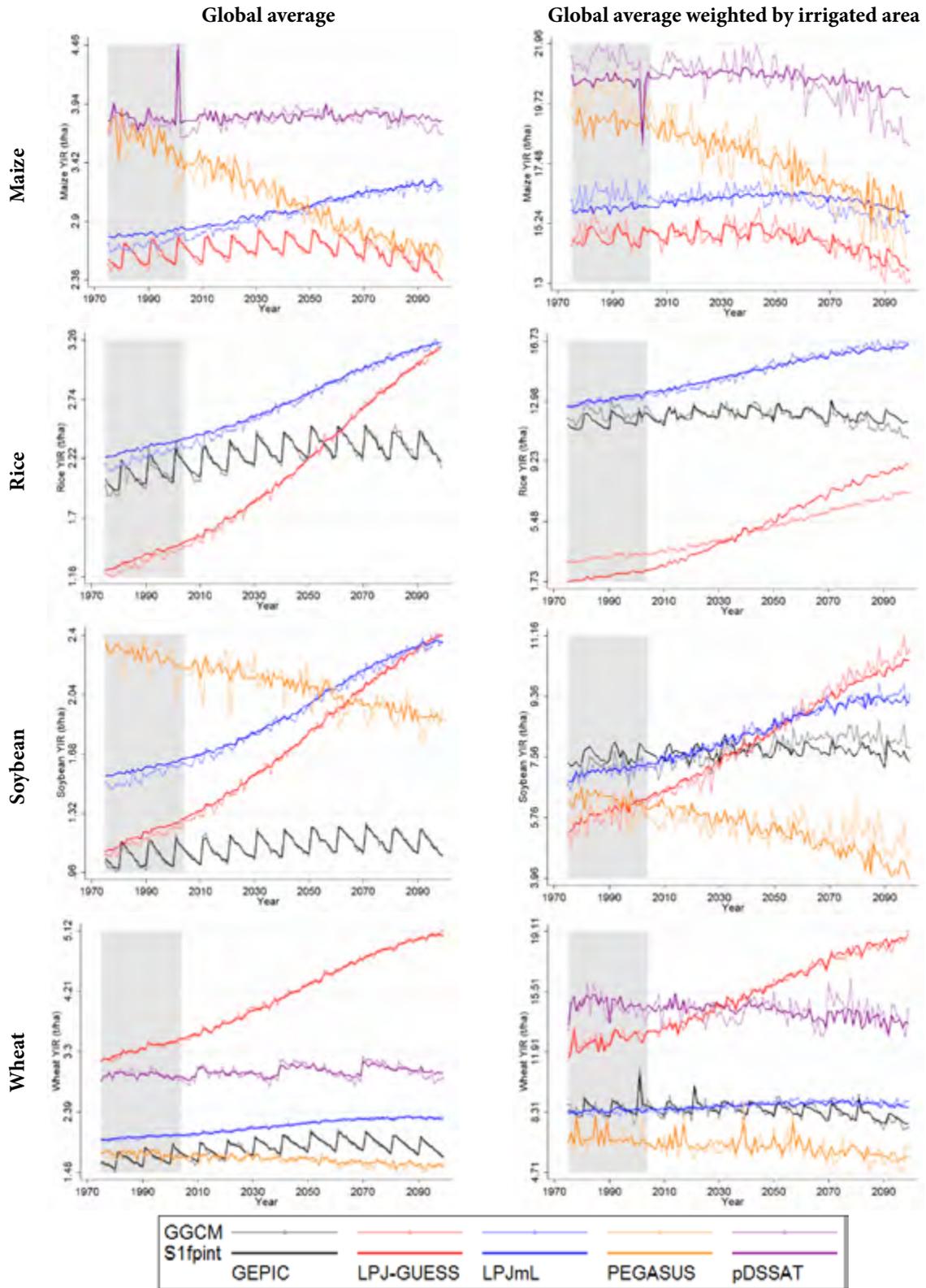
Figure 6. Average irrigated crop yields from GGCMs and statistical emulators (S1fpintsoil specification)

Note: Shaded areas represents the 'historical' period.

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.
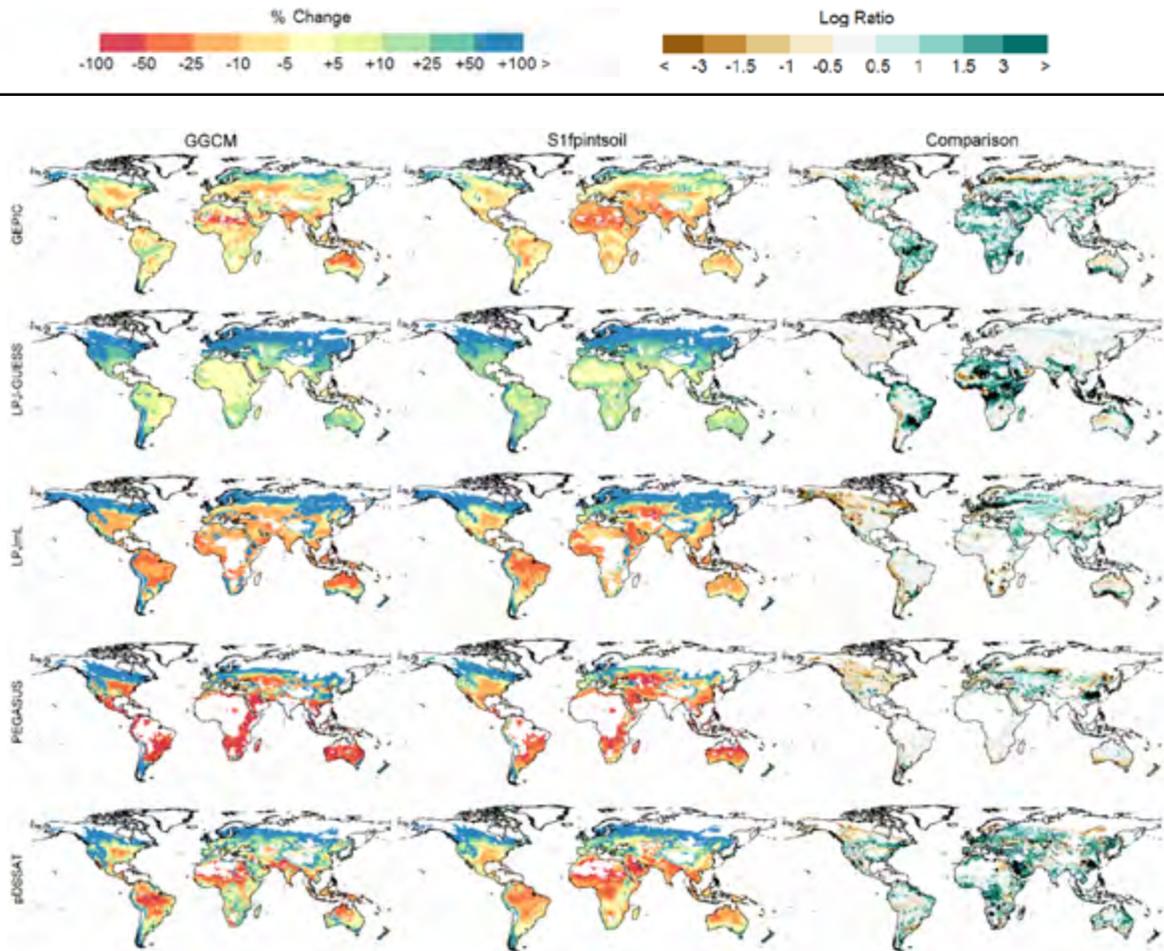
Figure 7. Changes in irrigated maize yields from 2000s to 2090s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio)
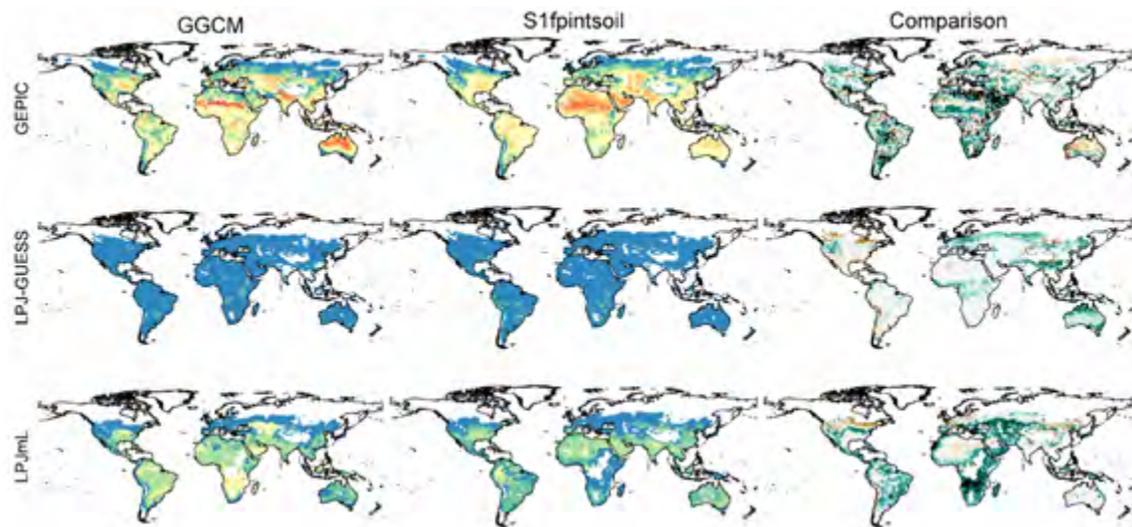
Figure 8. Changes in irrigated rice yields from 2000s to 2090s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio)
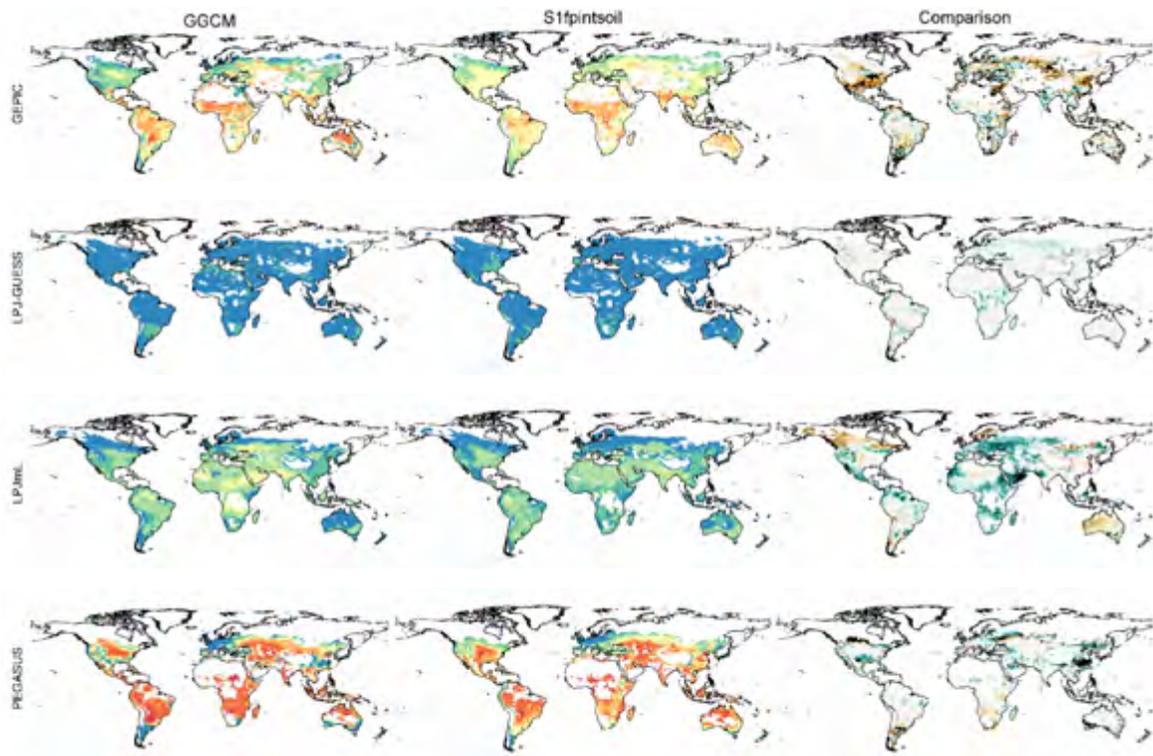
Figure 9. Changes in irrigated soybean yields from 2000s to 2090s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio)
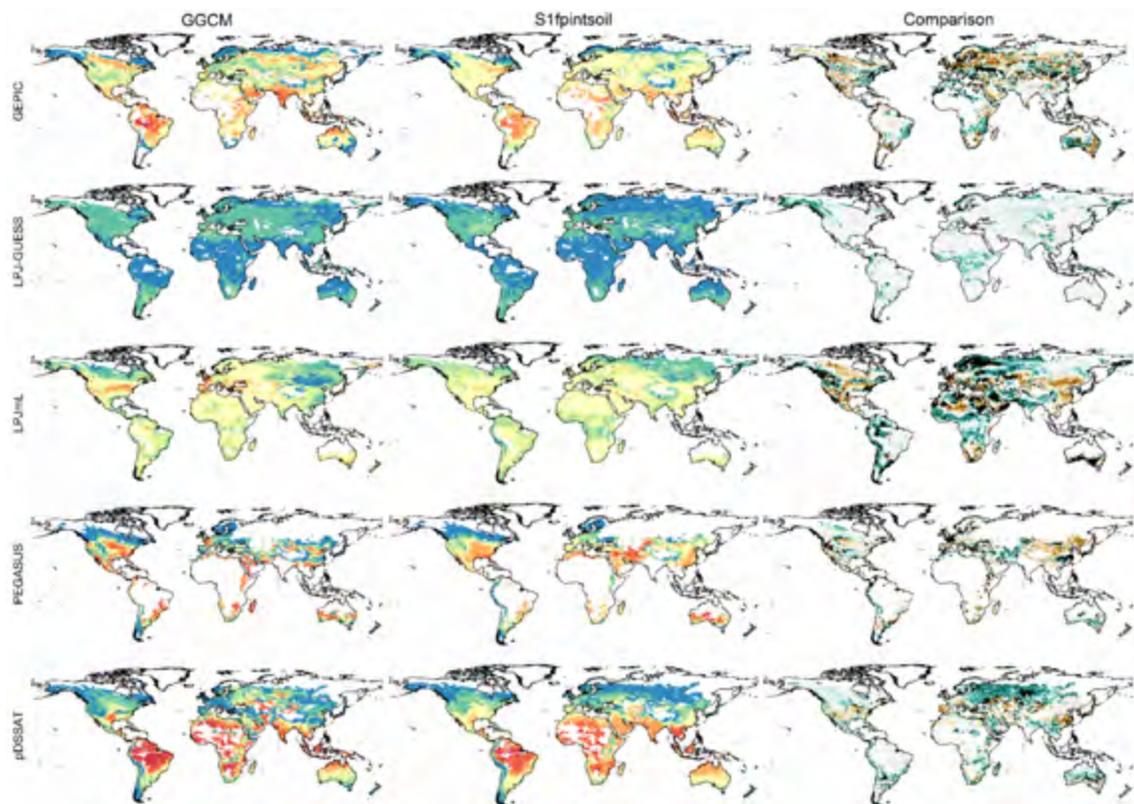
Figure 10. Changes in irrigated wheat yields from 2000s to 2090s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio)

by the end of the century. For other regions, the effects depend on the crop and model considered. However, the maps show that, overall, the emulators reproduce reasonably well the spatial patterns of climate change impacts on irrigated crop yields simulated by the GGCMs.

By the end of the century, substantial increases in irrigated maize yields are projected in the North and large decreases in the tropical regions by most GGCMs. Those patterns are reproduced by the emulators except for Northern America where yield increases are underestimated by the emulator for the LPJml, PEGASUS and pDSSAT models. Alternatively, the emulator overestimates irrigated maize yield increases in South America and Central Africa for the GEPIC and LPJ-GUESS models. Areas of disagreement regarding the sign of the impacts are observed in those regions for the LPJ-GUESS model.

For rice, generally good spatial agreement of yield changes is observed between the emulator and GGCMs projections for the LPJ-GUESS model. However, in accordance with the time series presented above, the emulator tends to overestimate irrigated yields in Southern Asia, where the largest part of irrigated rice is harvested. For the two other models, the pattern of agreement in this region is more mixed.

Irrigated soybean yield changes projected by the emulators are in high agreement with those from the LPJ-GUESS model in most regions. For the GEPIC model, however, the emulators tend to underestimate soybean yield changes in the southern part of the United States. For the PEGASUS model, disagreement regarding the sign of the impacts is observed in China, where a large share of soybean is harvested. However, Central US, which also support a large share of irrigated soybean is very well represented by the emulator for all models except GEPIC.

As for the other crops, the spatial patterns of changes in crop yields, projected by the GGCMs for wheat are also replicated by the emulators, especially the LPJ-GUESS model. However, the emulator of the GEPIC model tends to overestimate irrigated yields in eastern China where a large share of irrigated wheat is cultivated.

4.1.2 Irrigation water withdrawals

The same within-sample validation exercise as for irrigated crop yields is performed on irrigation water withdrawal estimates for each crop, grid cell, year, and climate model. **Figure 11** reports time series of irrigation water withdrawals averaged at the global level on the left and weighted by crop-specific irrigated harvested area on the right. The graphs show that projections from the emulator (in dark colors) follow the same trend as projections from GGCMs (in light colors). As for yields, inter-annual variability is emulated with less precision. However, divergences are

only observed for rice simulated by the LPJ-GUESS model when considering averages weighted by irrigated areas.

To assess the spatial agreement in irrigation water withdrawals estimated by the GGCMs and the statistical emulators (S1fpintsoil specification) for each crop, maps presenting climate change impact projections over the 2090s period are presented in Appendix H. These maps show that the emulators are able to reproduce the spatial patterns of irrigation demand over the globe for all crops. Maps representing spatial agreement in terms of changes from 2000s to 2090s are presented in **Figure 12** to **Figure 15**. The maps show that large decreases in irrigation demand are expected by most GGCMs. The patterns are reproduced reasonably well by the emulators, except for Northern Eurasia, where a large area of disagreement regarding the sign of the impact is observed for the GEPIC and PEGASUS models.

Regarding rice, changes in water withdrawals for the LPJmL model are expect to be largely positive, by contrast to the two other models. Those changes are emulated by the statistical model, with some overestimation by the emulator for the Northern Eurasia with the LPJ-GUESS model and in Asia for the GEPIC model. Contradictory impacts projections are observed in Central Africa and India for the LPJ-GUESS model.

For soybean, projected changes in irrigation water withdrawals are also largely positive for the LPJ-GUESS model with similar area of disagreement as those observed for rice. For PEGASUS, the decreases in water withdrawal for irrigation are projected to be very large by the end of century and the emulator represent those change very well.

Change in irrigation water withdrawals for wheat project by the GEPIC and LPJmL model are reproduced relatively well by the emulator. For the other models, larger regions or over- and underestimation are observed.

4.2 Out-of-sample validation

The out-of-sample validation exercise consists of comparing outputs from emulators which are re-estimated using a partial sample excluding simulations from one climate model, to outputs from GGCMs under the excluded climate model sub-sample. This exercise is performed for both irrigated yields and irrigation water withdrawal.

4.2.1 Irrigated crop yields

For irrigated crop yields, the NRMSE statistics calculated for each excluded sample are reported in **Table 5** and compared to the NRMSEs from the full sample estimation presented in Section 3. Unsurprisingly, the prediction errors from the out-of-sample exercise are larger than those from the in-sample estimations. The differences between the overall out-of-sample NRMSEs and the in-sample NRMSEs are however relatively small, with differences ranging between 0.002 and 0.009. The errors are generally the smallest for
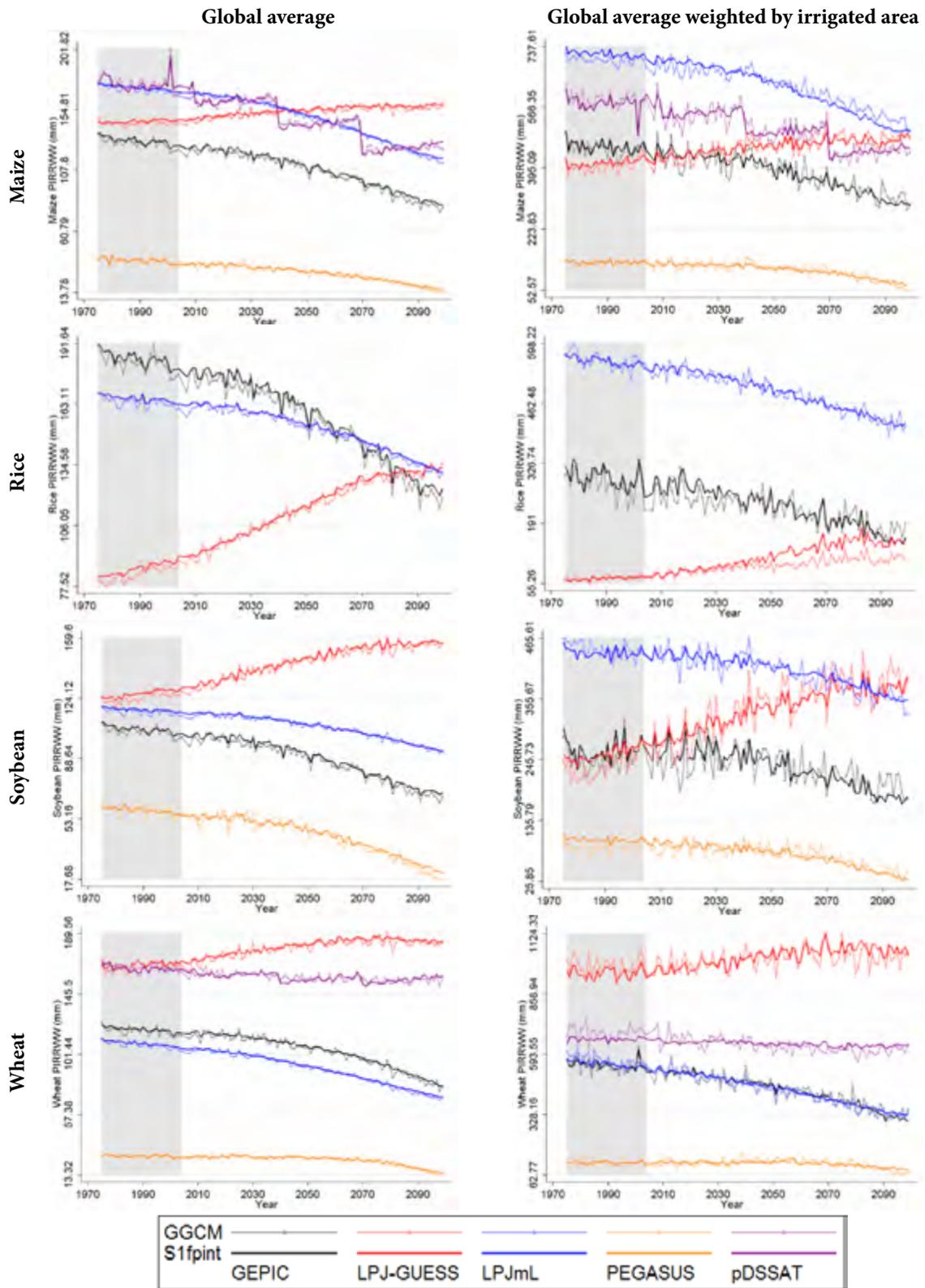
Figure 11. Average irrigation water withdrawals from GGCMs and statistical emulators (S1fpintsoil specification)

Note: Shaded areas represents the 'historical' period.

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.
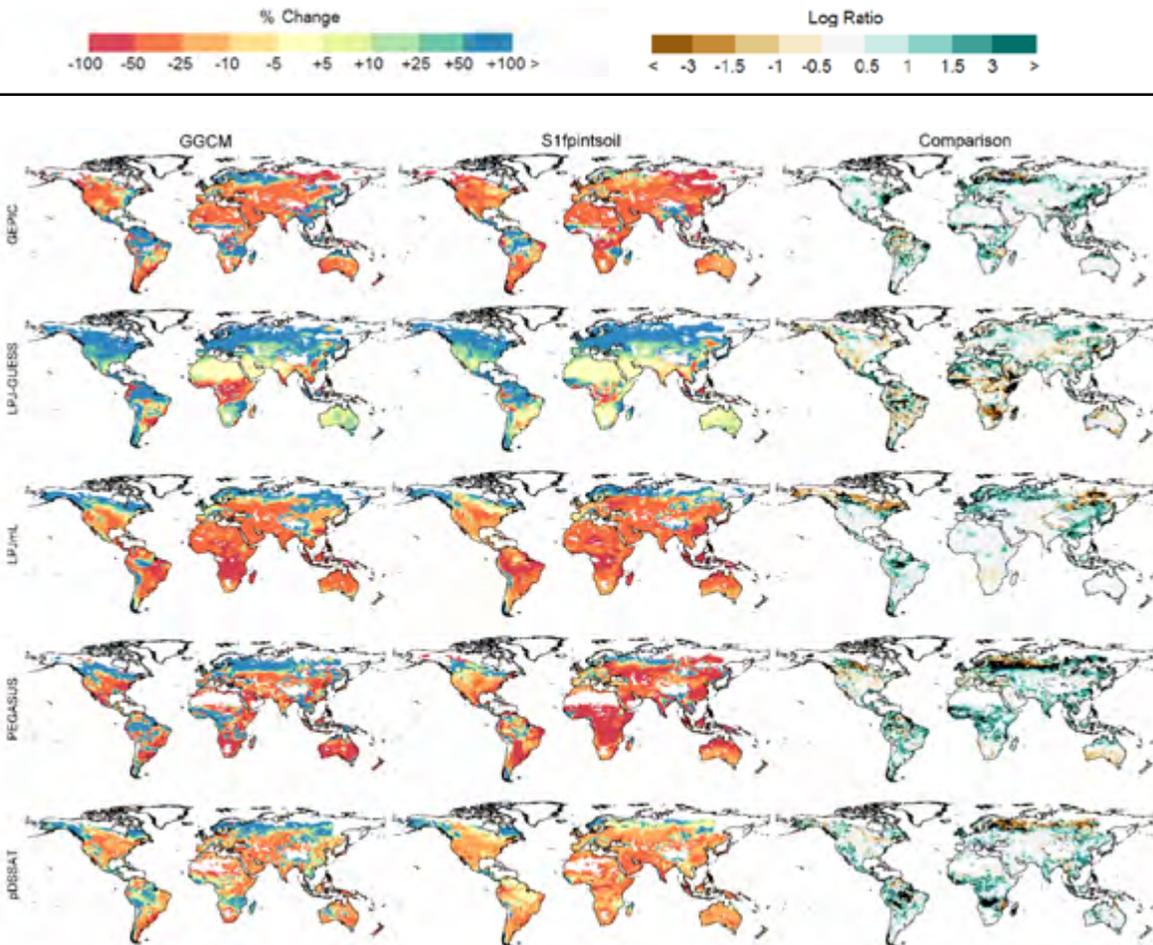

Figure 12. Changes in irrigation water withdrawals for maize from 2000s to 2090s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio)
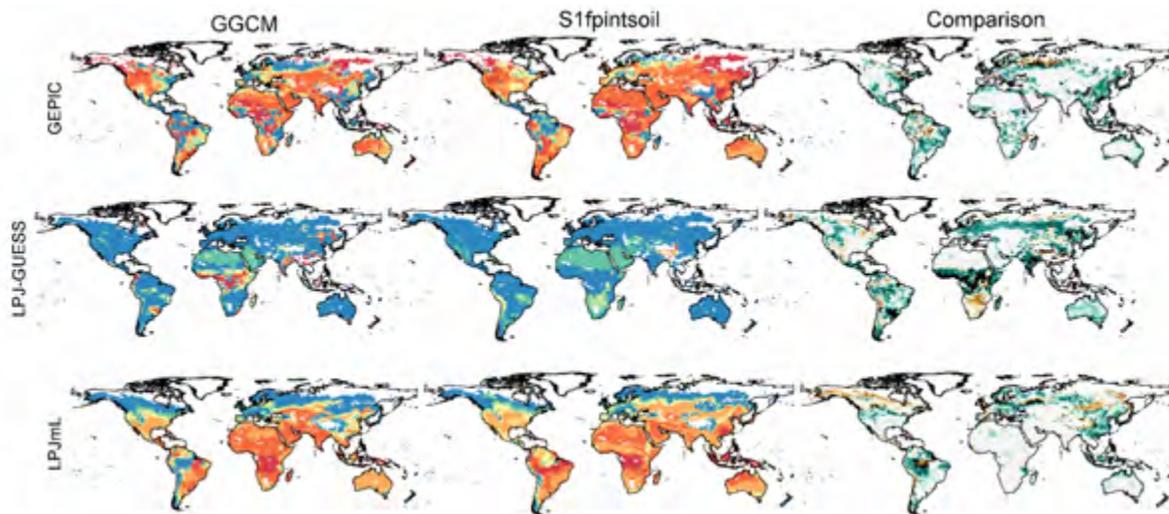
Figure 13. Changes in irrigation water withdrawals for rice from 2000s to 2090s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio)
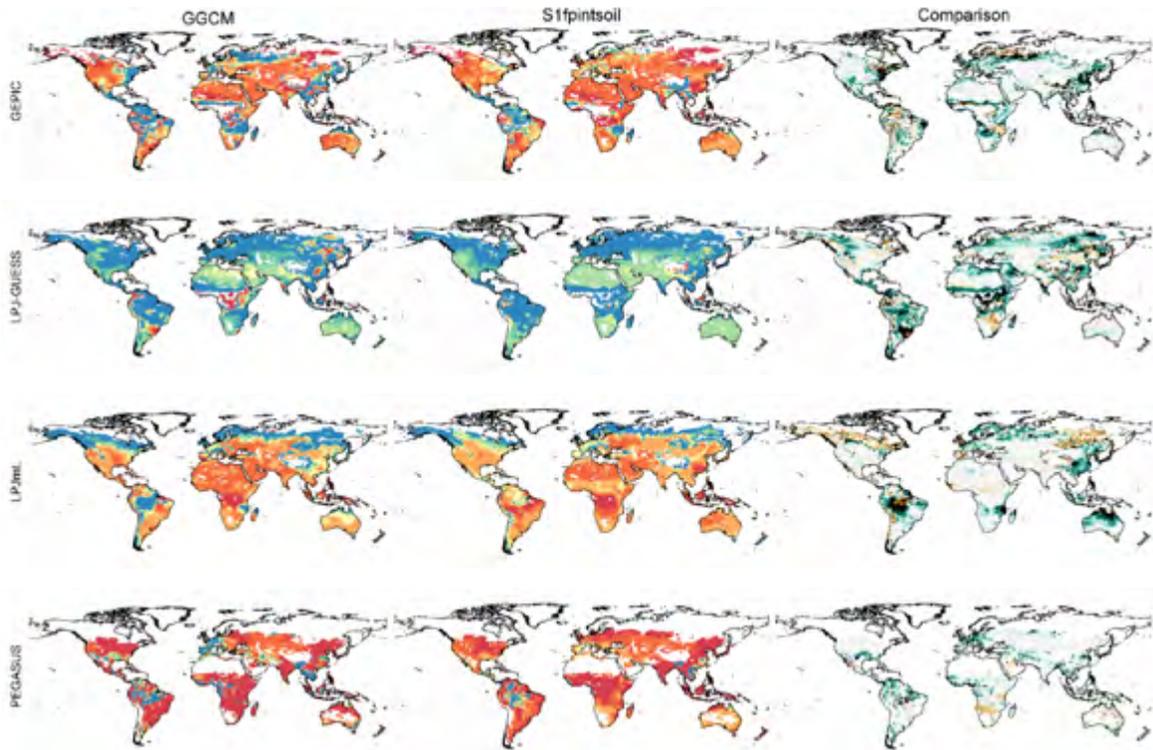
Figure 14. Changes in irrigation water withdrawals for soybean from 2000s to 2090s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio)
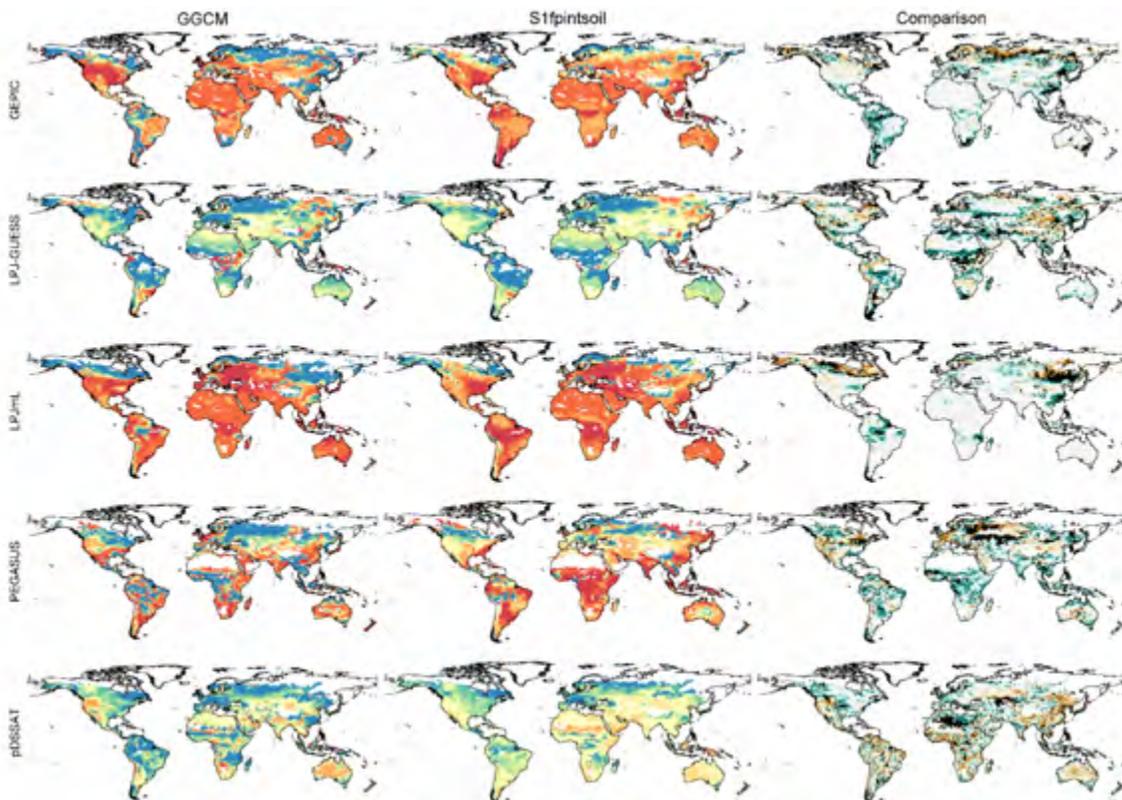
Figure 15. Changes in irrigation water withdrawals for wheat from 2000s to 2090s estimated by the statistical emulators (S1fpintsoil specification) and GGCMs and comparison (log ratio)

Table 5. NRMSE statistics for the leave-one-GCM-out validation (S1fpintsoil specification) compared to the full sample

Crop	Model	GFDL-ESM2M	HadGEM2-ES	NorESM1-M	Overall	Full sample
Maize	GEPIC	0.051	0.057	0.048	0.052	0.045
	LPJ-GUESS	0.049	0.045	0.042	0.045	0.037
	LPJmL	0.045	0.040	0.040	0.042	0.035
	pDSSAT	0.076	0.071	0.074	0.074	0.065
	PEGASUS	0.057	0.061	0.055	0.058	0.056
Rice	GEPIC	0.065	0.068	0.059	0.064	0.056
	LPJ-GUESS	0.044	0.038	0.038	0.040	0.033
	LPJmL	0.043	0.042	0.037	0.041	0.037
Soybean	GEPIC	0.066	0.066	0.054	0.062	0.054
	LPJ-GUESS	0.049	0.046	0.041	0.045	0.037
	LPJmL	0.035	0.034	0.031	0.033	0.030
	PEGASUS	0.048	0.052	0.043	0.048	0.042
Wheat	GEPIC	0.054	0.057	0.052	0.055	0.049
	LPJ-GUESS	0.039	0.037	0.036	0.038	0.029
	LPJmL	0.033	0.032	0.029	0.032	0.026
	pDSSAT	0.643	0.669	0.570	0.627	0.572
	PEGASUS	0.033	0.037	0.031	0.034	0.030

the estimates with the NorESM1-M climate model excluded from the estimation sample.

Time series of irrigated yield weighted by irrigated area harvested for each crop, GGCM and leave-one-GCM-out combination are presented in **Figure 16**. The graphs show that, as for the in-sample validation, the emulators are able to reproduce out-of-sample the trend in crop yields of most GGCMs. However, in some cases, the emulator and GGCM outputs differ depending on the climate sample excluded. For instance, for maize yields with the GEPIC model, the graphs indicate that, in the case where the HadGEM2-ES model is excluded from the training dataset, the emulated irrigated maize crop yields are overestimated while they are overestimated in the case where the NorESM1-M model is excluded. In such cases, the use of the largest sample of plausible climate change is essential to estimate the response functions.

4.2.2 Irrigation water withdrawals

As for irrigated crop yields, the NRMSE statistics calculated for irrigation water withdrawal for each excluded sample (see **Table 6**) show that the prediction errors from the out-of-sample exercise are slightly larger than those from the in-sample estimations, but the differences are larger than those observed for yields (differences ranging between 0.003 and 0.02) especially for the GEPIC model for maize and wheat. As for irrigated yields, the errors are generally the smallest under the excluded NorESM1-M climate model.

Time series of average irrigation water withdrawals weighted by irrigated area harvested are presented in **Figure 17** for each crop, GGCM and leave-one-GCM-out combination. The graphs show that out-of-sample *PIRRWW* are generally overestimated by the emulators in cases where projections from GGCMs are the smallest and underestimated where projections are the largest.

5. Conclusion

Based on the methodology developed in Blanc and Sultan (2015) and Blanc (2017), this analysis develops statistical emulators of global gridded crop models for irrigated crops yields and associated water withdrawals. The emulators for maize, rice, soybean and wheat are estimated using data from an ensemble of simulations from five GGCMs as part of the ISI-MIP Fast Track intercomparison exercise. Crop-specific response functions for each GGCM are estimated at the grid-cell level for both irrigated crop yields and irrigation water withdrawals.

To evaluate the statistical emulators' ability to reproduce irrigated crop yields and associated irrigation water withdrawals, both in-and out-of-sample validation exercises are conducted. These exercises show that, in most cases, outputs from the statistical emulators follow the same trend as projections from GGCMs. Inter-annual yield variability is captured with less accuracy but spatial analyses reveal that, overall, the emulators tend to capture the spatial patterns of climate change impacts on irrigated crop yields and irrigation water withdrawals. Similar spatial agreements
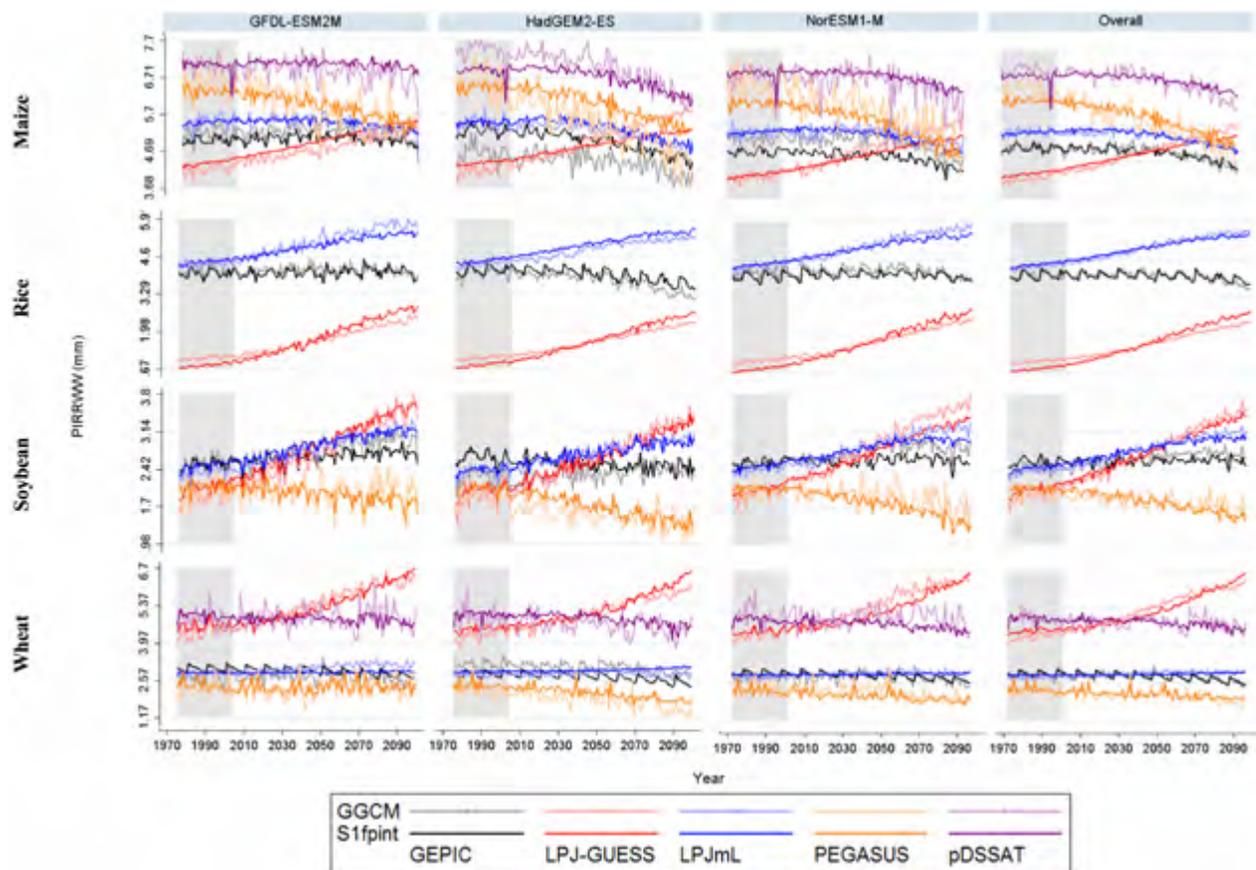
Figure 16. Average irrigated crop yield projections from GGCMs and statistical models (S1fpintsoil specification) weighted by irrigated area harvested in the leave-one-GCM-out validation exercise

Table 6. NRMSE statistics for the leave-one-GCM-out validation (S1fpintsoil specification) compared to the full sample

Crop	Model	GFDL-ESM2M	HadGEM2-ES	NorESM1-M	Overall	Full sample
Maize	GEPIC	0.070	0.081	0.078	0.076	0.061
	LPJ-GUESS	0.059	0.058	0.052	0.056	0.048
	LPJmL	0.049	0.045	0.042	0.045	0.039
	pDSSAT	0.084	0.074	0.074	0.077	0.067
	PEGASUS	0.044	0.047	0.040	0.044	0.040
Rice	GEPIC	0.061	0.071	0.063	0.065	0.054
	LPJ-GUESS	0.048	0.049	0.043	0.047	0.038
	LPJmL	0.054	0.048	0.045	0.049	0.041
Soybean	GEPIC	0.068	0.075	0.068	0.070	0.060
	LPJ-GUESS	0.055	0.061	0.051	0.056	0.045
	LPJmL	0.055	0.049	0.048	0.051	0.043
	PEGASUS	0.056	0.057	0.049	0.054	0.049
Wheat	GEPIC	0.096	0.103	0.094	0.098	0.078
	LPJ-GUESS	0.060	0.059	0.052	0.057	0.048
	LPJmL	0.053	0.053	0.050	0.052	0.042
	pDSSAT	0.055	0.045	0.050	0.050	0.030
	PEGASUS	0.043	0.043	0.036	0.041	0.035
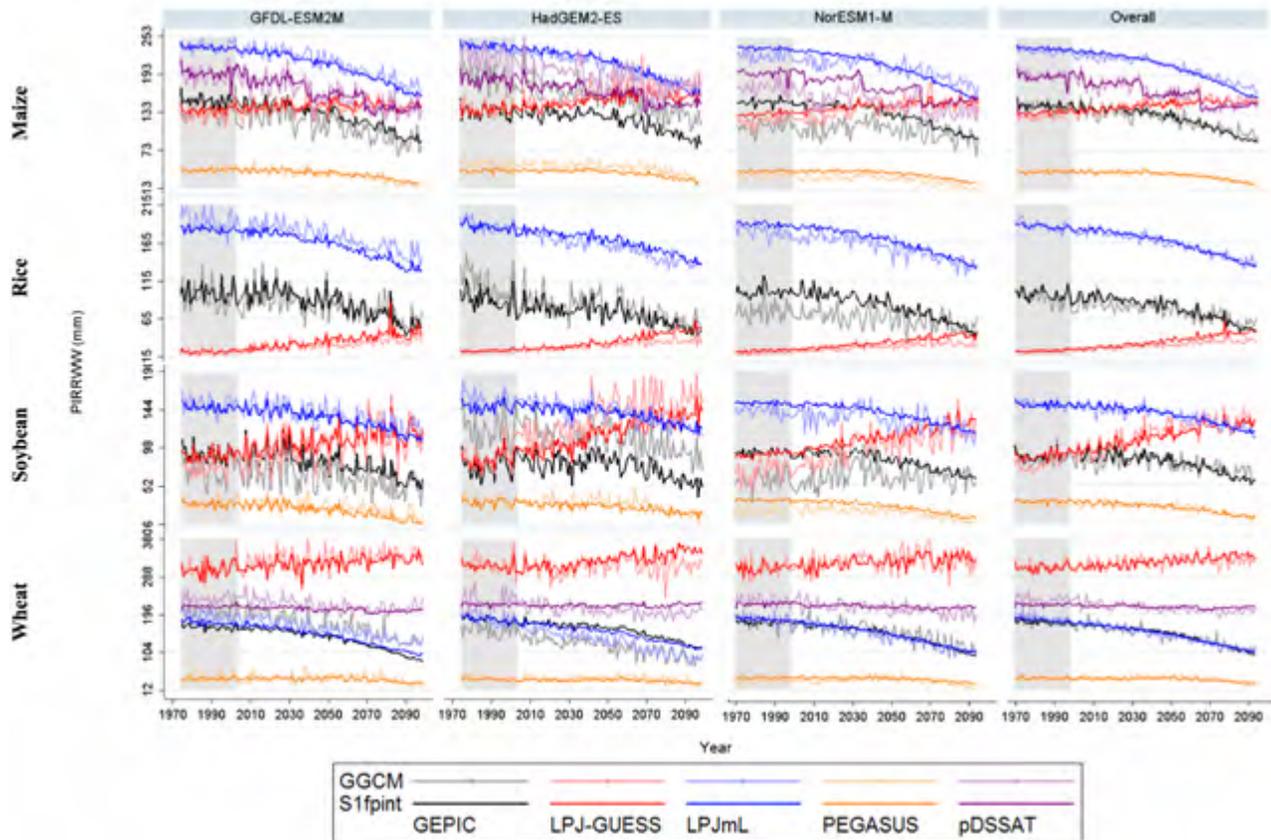
Figure 17. Average irrigation water withdrawal projections from GGCMs and statistical models (S1fpintsoil specification) weighted by irrigated area harvested in the leave-one-GCM-out validation exercise

are observed when considering the changes in outputs between the beginning and end of the century, despite some disagreements regarding the strength of the impacts in different regions depending on the GGCM considered. When using the emulators for regional assessments of climate change impacts, caution should therefore be exercised when selecting an ensemble of emulators that best capture the impact projected by the underlying GGCMs.

Out-of-sample validation exercises also show a general agreement between the emulators and the GGCMs. However, as expected, prediction accuracy is lowered when excluding output responses to weather variables outside the range of values found in the estimation sample. Estimating the statistical emulator using the largest sample available, which is designed to encompass the largest range of plausible changes in climate over the century, is essential.

The statistical emulators estimated in this study offer an accessible and reliable tool to estimate climate change impacts on irrigated crop yields and associated irrigation water withdrawals under alternative plausible user-defined

scenarios. However, as previously noted in Blanc (2017), the emulator is better suited to assess long-term climate change impacts rather than inter-annual. It is also important to note that as none of the GGCMSs is considered more accurate than another at projecting future crop yields, predictions from multiple models should be considered. In this regard, this study developed the emulators for several crop models to provide a computationally efficient way to consider modeling uncertainty in climate change impact assessments.

Acknowledgments

6. References

Auffhammer, M., and W. Schlenker (2014): Empirical Studies on Agricultural Impacts and Adaptation. *Energy Economics* 46 (November): 555–61. (doi:10.1016/J.ENECO.2014.09.010).

Blanc, É. (2017): Statistical Emulators of Maize, Rice, Soybean and Wheat Yields from Global Gridded Crop Models. *Agricultural and Forest Meteorology* 236: 145–61.

Blanc, É., J. Caron, C. Fant, and E. Monier (2017): Climate Change Impact on Water Resources and Irrigated Crop Yields in the US. *Earth's Future*, 5(8): 877–892.

Blanc, É., and E. Strobl (2013): The Impact of Climate Change on Cropland Productivity: Evidence from Satellite Based Products at the River Basin Scale in Africa. *Climatic Change* 117(4): 873–90.

Blanc, É., and B. Sultan (2015): Emulating Maize Yields from Global Gridded Crop Models Using Statistical Estimates. *Agricultural and Forest Meteorology* 214–215: 134–47. (doi:10.1016/j.agrformet.2015.08.256).

Bondeau, A., P.C. Smith, S. Zaehle, S. Schaphoff, W. Lucht, W. Cramer, D. Gerten, *et al.* (2007): Modelling the Role of Agriculture for the 20th Century Global Terrestrial Carbon Balance. *Global Change Biology* 13(3): 679–706. (doi:10.1111/j.1365-2486.2006.01305.x).

Boote, K.J., J.W. Jones, J.W. White, S. Asseng, and J.I. Lizaso (2013): Putting Mechanisms into Crop Production Models. *Plant, Cell & Environment* 36(9): 1658–72. (doi:10.1111/pce.12119).

Calvin, K., and K. Fisher-Vanden (2017): Quantifying the Indirect Impacts of Climate on Agriculture: An Inter-Method Comparison. *Environmental Research Letters* 12(11): 115004. (doi:10.1088/1748-9326/aa843c).

Deryng, D., D. Conway, N. Ramankutty, J. Price, and R. Warren (2014): Global Crop Yield Response to Extreme Heat Stress under Multiple Climate Change Futures. *Environmental Research Letters* 9(3): 34011. http://stacks.iop.org/1748-9326/9/i=3/a=034011.

Deryng, D., W.J. Sacks, C.C. Barford, and N. Ramankutty (2011): Simulating the Effects of Climate and Agricultural Management Practices on Global Crop Yield. *Global Biogeochemical Cycles* 25(2). (doi:10.1029/2009GB003765).

Elliott, J., M. Glotter, N. Best, D. Kelly, M. Wilde, and I. Foster (2013): The Parallel System for Integrating Impact Models and Sectors (PSIMS). In *Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery (XSEDE '13)*, 21: 1–8. Association for Computing Machinery.

Elliott, J., D. Deryng, C. Müller, K. Frieler, M. Konzmann, D. Gerten, M. Glotter, *et al.* (2014): Constraints and Potentials of Future Irrigation Water Availability on Agricultural Production under Climate Change. *PNAS* 111(9). (doi:10.1073/pnas.1222474110).

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

Haim, D., M. Shechter, and P. Berliner (2007): Assessing the Impact of Climate Change on Representative Field Crops in Israeli Agriculture: A Case Study of Wheat and Cotton. *Climatic Change* 86(3–4): 425–440

Hempel, S., K. Frieler, L. Warszawski, J. Schewe, and F. Piontek (2013): A Trend-Preserving Bias Correction – the ISI-MIP Approach. *Earth System Dynamics* 4(2): 219–36.

Holzkämper, A., P. Calanca, and J. Fuhrer (2012): Statistical Crop Models: Predicting the Effects of Temperature and Precipitation Changes. *Climate Research* 51(1): 11–21. (doi:10.3354/cr01057).

Hsiang, S. (2016): Climate Econometrics. *Annual Review of Resource Economics* 8(1): 43–75. (doi:10.1146/annurev-resource-100815-095343).

Jones, J., G. Hoogenboom, C. Porter, K. Boote, W. Batchelor, L. Hunt, and J. Ritchie (2003): The DSSAT Cropping System Model. *European Journal of Agronomy* 18(3–4): 235–65.

Lindeskog, M., A. Arneth, A. Bondeau, K. Waha, J. Seaquist, S. Olin, and B. Smith (2013): Implications of Accounting for Land Use in Simulations of Ecosystem Services and Carbon Cycling in Africa. *Earth System Dynamics Discussions* 4: 235–78.

Liu, J., J.R. Williams, A.J.B. Zehnder, and H. Yang (2007): GEPIC – Modelling Wheat Yield and Crop Water Productivity with High Resolution on a Global Scale. *Agricultural Systems* 94(2): 478–93. (doi:10.1016/j.agsy.2006.11.019).

Lobell, D.B., and S. Asseng (2017): Comparing Estimates of Climate Change Impacts from Process-Based and Statistical Crop Models. *Environmental Research Letters* 12(1): 15001.

Lobell, D.B., and M.B. Burke (2010): On the Use of Statistical Models to Predict Crop Yield Responses to Climate Change. *Agricultural and Forest Meteorology* 150(11): 1443–52. (doi:10.1016/j.agrformet.2010.07.008).

Lobell, D., and C. Field (2007): Global Scale Climate-Crop Yield Relationships and the Impacts of Recent Warming. *Environmental Research Letters* 2(1): 1–7.

Monier, E., S. Paltsev, A. Sokolov, Y.-H.H. Chen, X. Gao, Q. Ejaz, E. Couzo, *et al.* (2018): Toward a Consistent Modeling Framework to Assess Multi-Sectoral Climate Impacts. *Nature Communications* 9(1): 660. (doi:10.1038/s41467-018-02984-9).

Moore, F.C., U.L.C. Baldos, and T. Hertel (2017): Economic Impacts of Climate Change on Agriculture: A Comparison of Process-Based and Statistical Yield Models. *Environmental Research Letters* 12(6): 65008. http://stacks.iop.org/1748-9326/12/i=6/a=065008.

Ostberg, S., J. Schewe, K. Childers, and K. Frieler (2018): Changes in Crop Yields and Their Variability at Different Levels of Global Warming. *Earth System Dynamics Discussions* 9: 479–96. (doi:10.5194/esd-2017-69).

Oyebamiji, O.K., N.R. Edwards, P.B. Holden, P.H. Garthwaite, S. Schaphoff, and D. Gerten (2015): Emulating Global Climate Change Impacts on Crop Yields. *Statistical Modelling*, 15(6): 499–525. (doi:10.1177/1471082X14568248).

Parry, M., G. Fischer, M. Livermore, C. Rosenzweig, and A. Iglesias (1999): Climate Change and World Food Security: A New Assessment. *Global Environmental Change* 9: 551–67.

Portmann, F.T., S. Siebert, and P. Döll (2010): MIRCA2000 - Global Monthly Irrigated and Rainfed Crop Areas around the Year 2000: A New High-Resolution Data Set for Agricultural and Hydrological Modeling. *Global Biogeochemical Cycles* 24.

Riahi, K., A. Grübler, and N. Nakicenovic (2007): Scenarios of Long-Term Socio-Economic and Environmental Development under Climate Stabilization. *Technological Forecasting and Social Change* 74(7): 887–935. (doi:10.1016/j.techfore.2006.05.026).

Roberts, M.J., N.O. Braun, T.R. Sinclair, D.B. Lobell, and W. Schlenker (2017): Comparing and Combining Process-Based Crop Models and Statistical Models with Some Implications for Climate Change. *Environmental Research Letters* 12(9): 095010. (doi:10.1088/1748-9326/aa7f33).

Rosenzweig, C., J. Elliott, D. Deryng, A.C. Ruane, C. Müller, A. Arneth, K.J. Boote, *et al.* (2014): Assessing Agricultural Risks of Climate Change in the 21st Century in a Global Gridded Crop Model Intercomparison. *PNAS* 111(9): 3268–73. (doi:10.1073/pnas.1222463110).

Rosenzweig, C., and M. Parry (1994a): Potential Impacts of Climate Change on World Food Supply. *Nature* 367: 133–38.

Rosenzweig, C., and M.L. Parry (1994b): Potential Impact of Climate Change on World Food Supply. *Nature* 367(6459): 133–38. (doi:10.1038/367133a0).

Ruane, A.C., C. Rosenzweig, S. Asseng, K.J. Boote, J. Elliott, F. Ewert, J.W. Jones, *et al.* (2017): An AgMIP Framework for Improved Agricultural Representation in Integrated Assessment Models. *Environmental Research Letters* 12(12): 125003. http://stacks.iop.org/1748-9326/12/i=12/a=125003.

Schauberger, B., S. Archontoulis, A. Arneth, J. Balkovic, P. Ciais, D. Deryng, J. Elliott, *et al.* (2017): Consistent Negative Response of US Crops to High Temperatures in Observations and Crop Models. *Nature Communications* 8 (January): 13931.

Schlenker, W., and M.J. Roberts (2009): Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields under Climate Change. *PNAS* 106(37): 15594–98. (doi:10.1073/pnas.0906865106).

Smith, B., I.C. Prentice, and M.T. Sykes (2001): Representation of Vegetation Dynamics in the Modelling of Terrestrial Ecosystems: Comparing Two Contrasting Approaches within European Climate Space. *Global Ecology and Biogeography* 10(6): 621–37. (doi:10.1046/j.1466-822X.2001.t01-1-00256.x).

Taylor, K.E., R.J. Stouffer, and G.A. Meehl (2012): An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society* 93(4): 485–98. (doi:10.1175/BAMS-D-11-00094.1).

Waha, K., L.G.J. van Bussel, C. Müller, and A. Bondeau (2012): Climate-Driven Simulation of Global Crop Sowing Dates. *Global Ecology and Biogeography* 21(2): 247–59. (doi:10.1111/j.1466-8238.2011.00678.x).

Warszawski, L., K. Frieler, V. Huber, F. Piontek, O. Serdeczny, and J. Schewe (2014): The Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP): Project Framework. *PNAS* 111(9): 3228–32. (doi:10.1073/pnas.1312330110).

White, J.W., G. Hoogenboom, B.A. Kimball, and G.W. Wall (2011): Methodologies for Simulating Impacts of Climate Change on Crop Production. *Field Crops Research* 124(3): 357–68. (doi:10.1016/J.FCR.2011.07.001).

Williams, J.R., and V.P. Singh (1995): Chapter 25. The EPIC. In: *Computer Models of Watershed Hydrology*, 909–1000. Littleton, CO: Water Resources Publications.

Appendices

Appendix A. Data information

Table A1. Modeling group information

Model	Institution	Modelers' names
GEPIC	EAWAG (Switzerland)	Christian Folberth
LPJ-GUESS	Institutionen för naturgeografi och ekosystemvetenskap (INES), Lunds Universitet (Sweden)	Thomas Pugh, Stephan Olin
LPJmL	PIK (Germany)	Christoph Muller
PEGASUS	Tyndall Centre, University of East Anglia (UK)	Delphine Deryng
pDSSAT	University of Chicago (USA)	Joshua Elliott
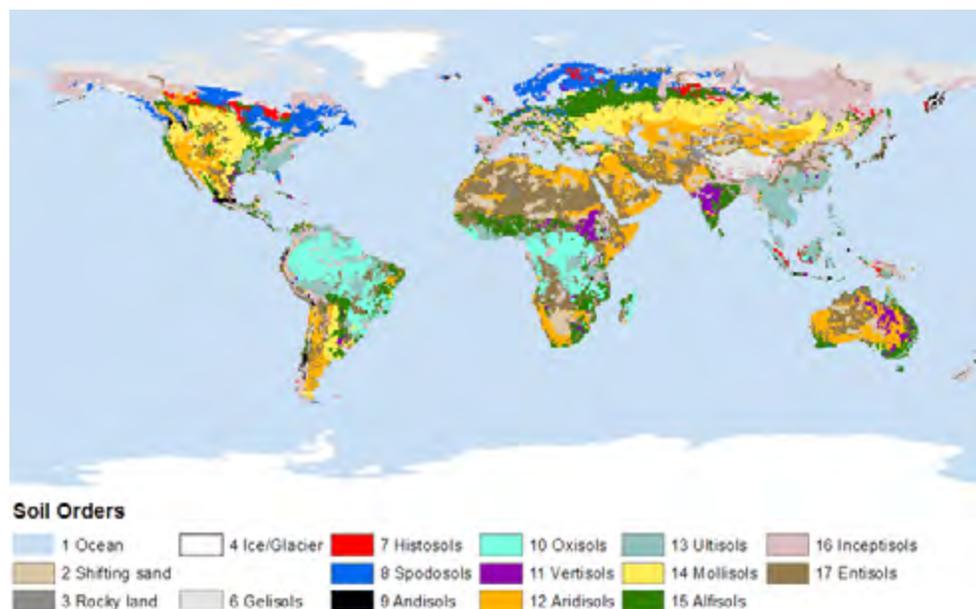
Figure A1. Global soil regions based on the FAO-UNESCO Soil Map of the World using the USDA soil taxonomy

Appendix B. Fractional polynomial transformation

See Excel file *Appendix_B_Variable_transformations.xslx* attached composed of the following table:

Table B1. Variable formulas for fractional polynomial transformation used in specification S1fpintsoil for *YIR*

Table B2. Variable formulas for fractional polynomial transformation used in specification S1fpintsoil for *PIRWW*

Appendix C. Regression results for *YIR* (S1fpintsoil specification)

See Excel file *Appendix_C_regression_results_YIR.xls* attached composed of the following tables:

Table C1. Regression results for maize *YIR* at the soil order level (specification S1fpintsoil)

Table C2. Regression results for rice *YIR* at the soil order level (specification S1fpintsoil)

Table C3. Regression results for soybean *YIR* at the soil order level (specification S1fpintsoil)

Table C4. Regression results for wheat *YIR* at the soil order level (specification S1fpintsoil)

Appendix D. Regression results for *PIRRWW* (S1fpintsoil specification)

See Excel file *Appendix_D_regression_results_PIRRWW.xls* attached composed of the following tables:

Table D1. Regression results for maize *PIRRWW* at the soil order level (specification S1fpintsoil)

Table D2. Regression results for rice *PIRRWW* at the soil order level (specification S1fpintsoil)

Table D3. Regression results for soybean *PIRRWW* at the soil order level (specification S1fpintsoil)

Table D4. Regression results for wheat *PIRRWW* at the soil order level (specification S1fpintsoil)

Appendix E. Fixed effects (δ) for *YIR* (S1fpintsoil specification)

See Excel file *Appendix_E_Grid_cells_FE_yir.xls* attached composed of the following tables:

Table E1. Grid cell fixed effect (δ) by GGCM for maize

Table E2. Grid cell fixed effect (δ) by GGCM for rice

Table E3. Grid cell fixed effect (δ) by GGCM for soybean

Table E4. Grid cell fixed effect (δ) by GGCM for wheat

Appendix F. Fixed effects (δ) for *PIRRWW* (S1fpintsoil specification)

See Excel file *Appendix_F_Grid_cells_FE_pirrww.xls* attached composed of the following tables:

Table F1. Grid cell fixed effect (δ) by GGCM for maize

Table F2. Grid cell fixed effect (δ) by GGCM for rice

Table F3. Grid cell fixed effect (δ) by GGCM for soybean

Table F4. Grid cell fixed effect (δ) by GGCM for wheat

Appendix G. In-sample validation for *YIR* (S1fpintsoil specification)
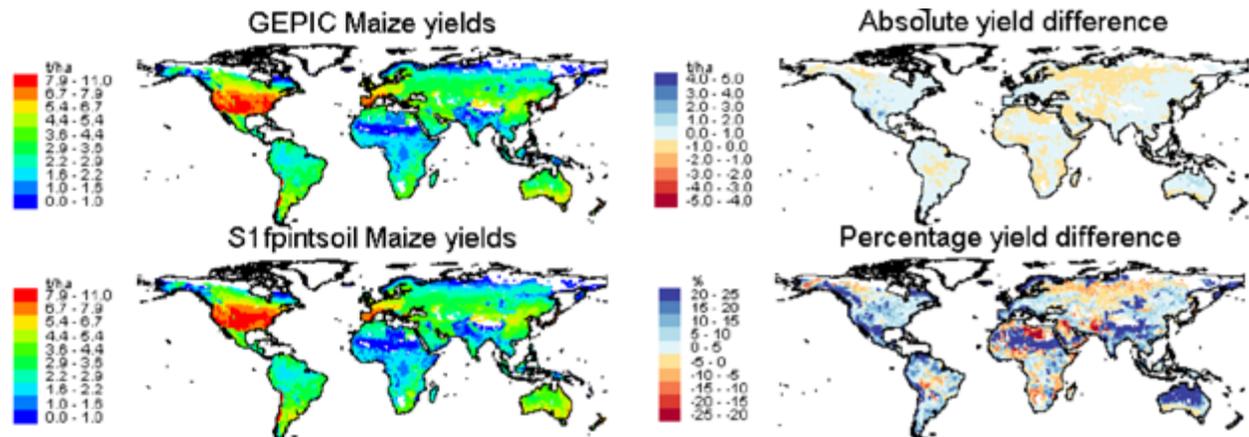
Figure G1. Irrigated maize yields averaged over 2090–2099 for the GEPIC model and S1fpintsoil specification
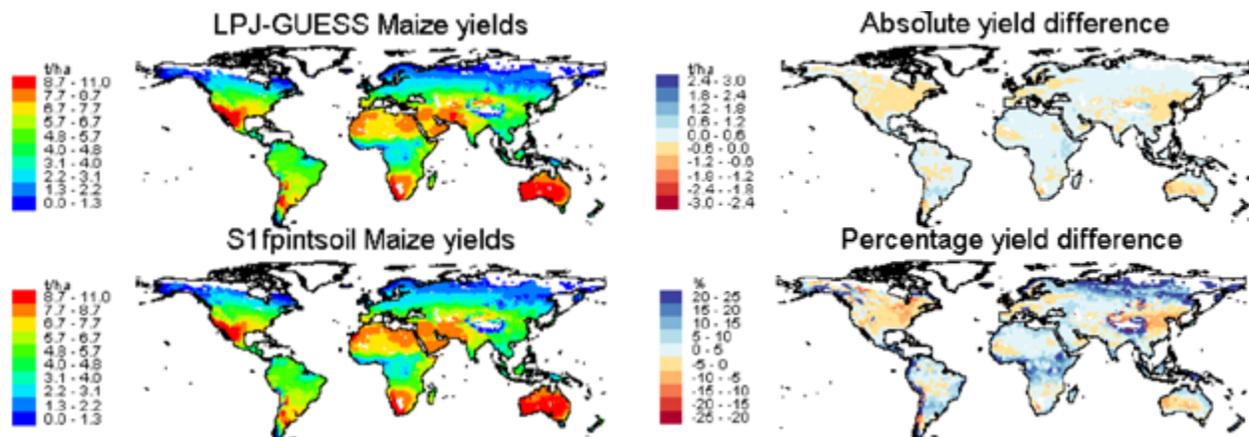
Figure G2. Irrigated maize yields averaged over 2090–2099 for the LPJ-GUESS model and S1fpintsoil specification
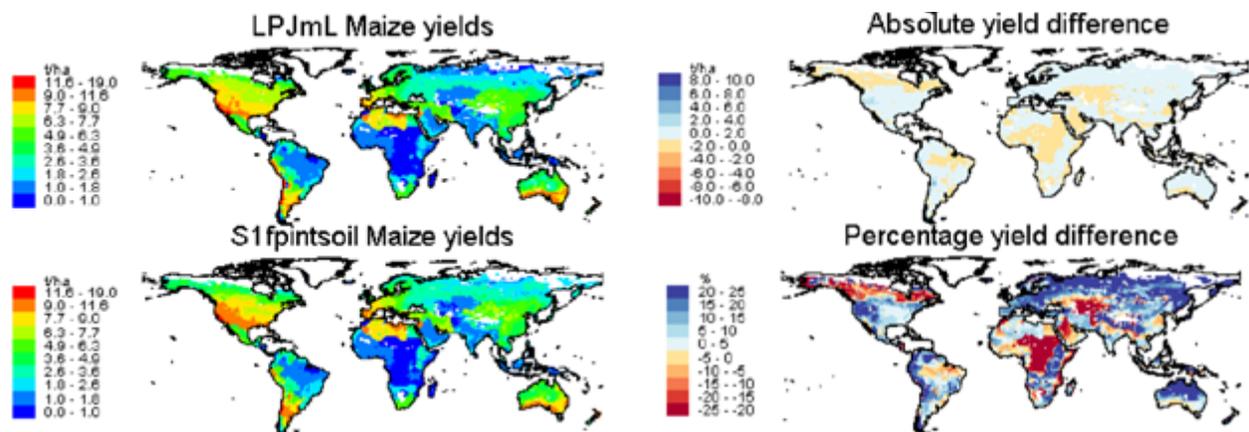
**Figure G3.** Irrigated maize yields averaged over 2090–2099 for the LPJmL model and S1fpintsoil specification
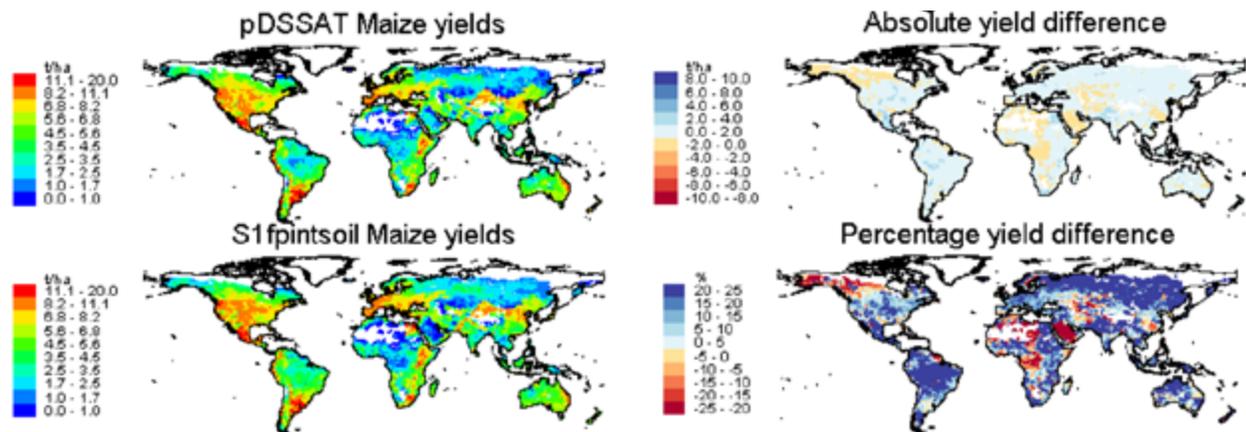
**Figure G4**. Irrigated maize yields averaged over 2090–2099 for the pDSSAT model and S1fpintsoil specification
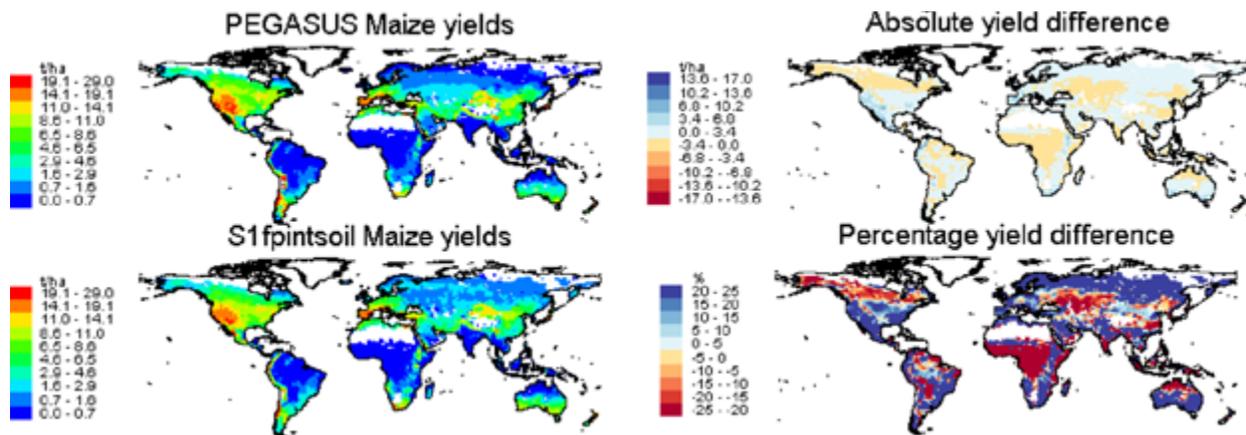
Figure G5. Irrigated maize yields averaged over 2090–2099 for the PEGASUS model and S1fpintsoil specification
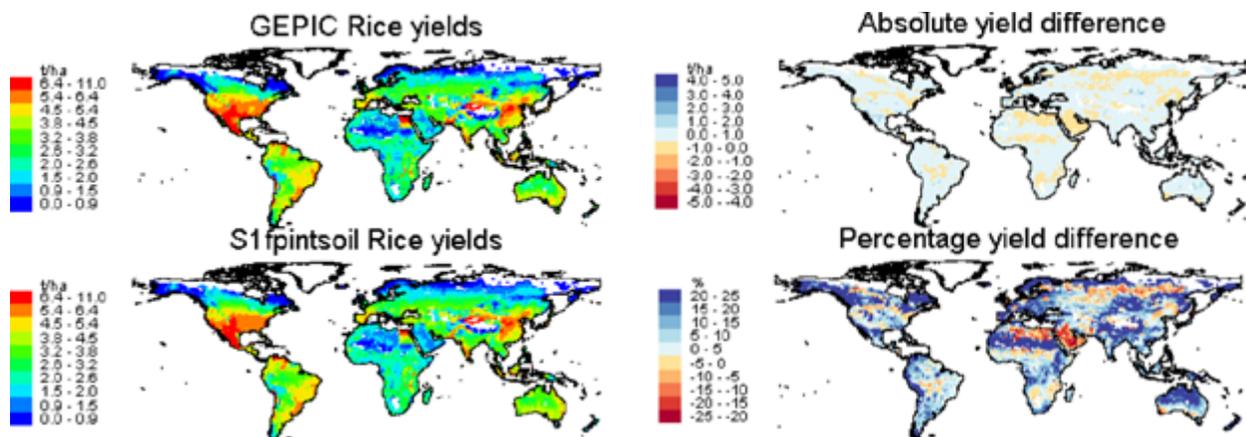
Figure G6. Irrigated rice yields averaged over 2090–2099 for the GEPIC model and S1fpintsoil specification
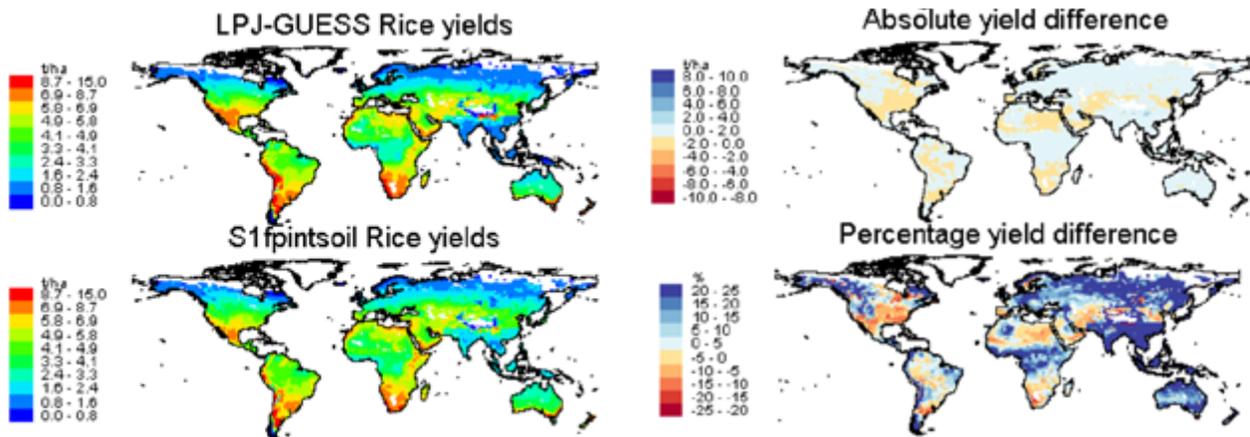
Figure G7. Irrigated rice yields averaged over 2090–2099 for the LPJ-GUESS model and S1fpintsoil specification
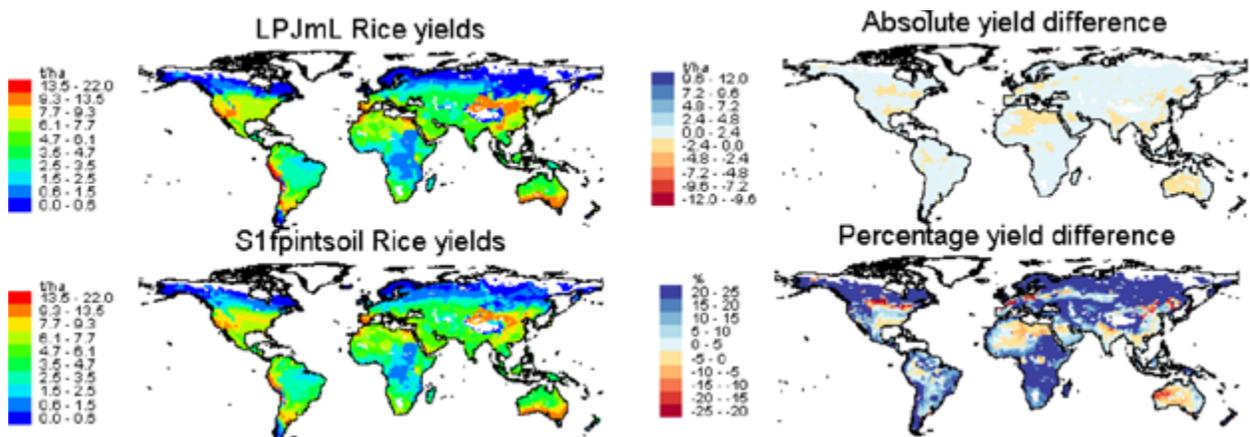
Figure G8. Irrigated rice yields averaged over 2090–2099 for the LPJmL model and S1fpintsoil specification
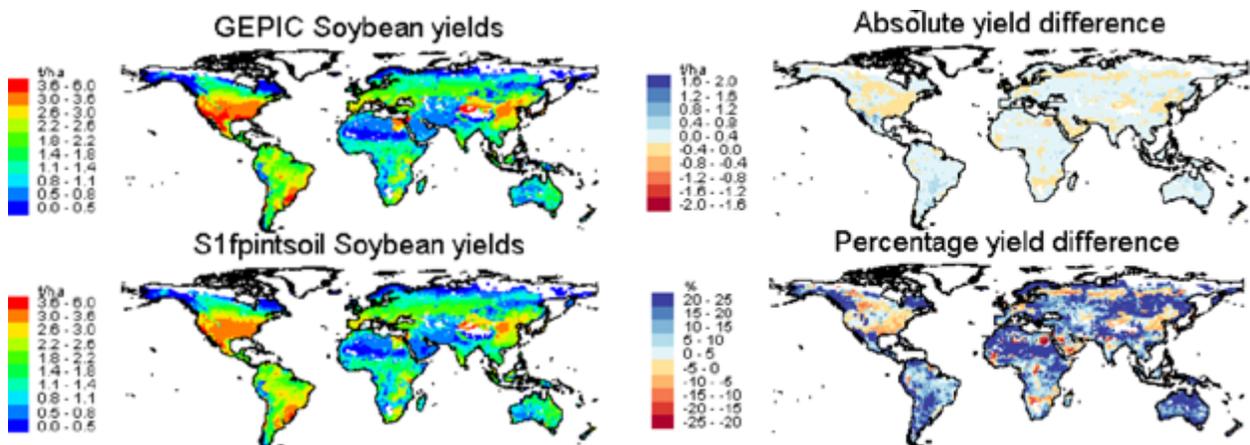
Figure G9. Irrigated soybean yields averaged over 2090–2099 for the GEPIC model and S1fpintsoil specification
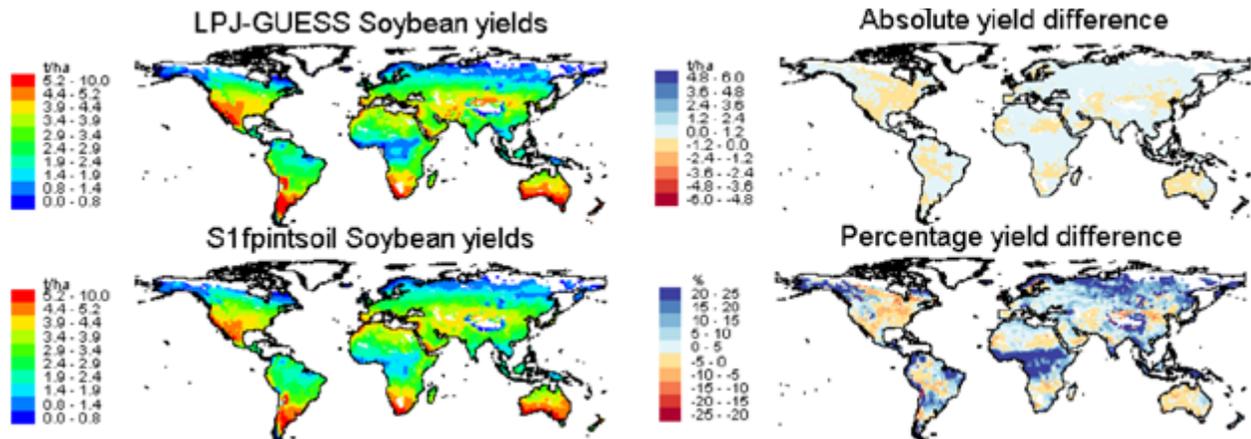
Figure G10. Irrigated soybean yields averaged over 2090–2099 for the LPJ-GUESS model and S1fpintsoil specification
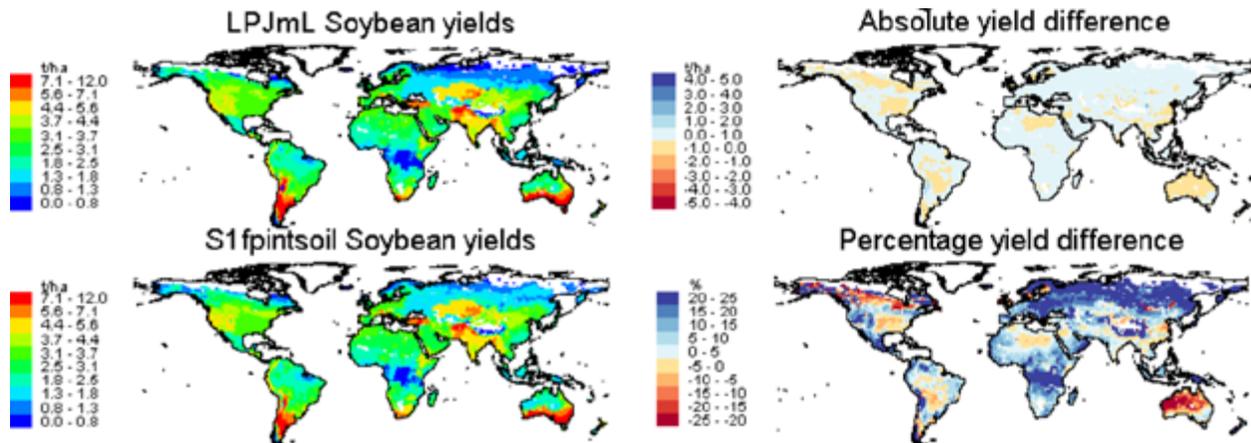
Figure G11. Irrigated soybean yields averaged over 2090–2099 for the LPJmL model and S1fpintsoil specification
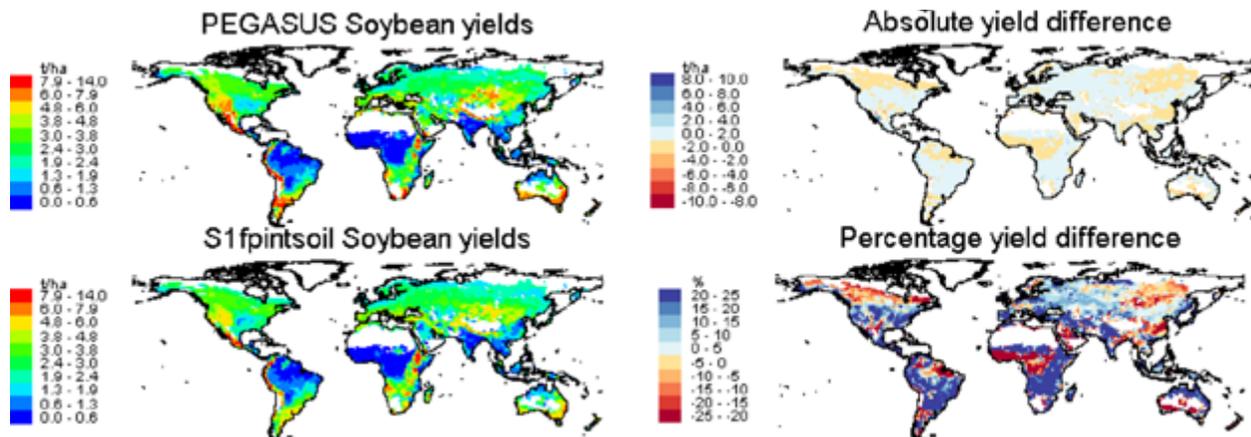
Figure G12. Irrigated soybean yields averaged over 2090–2099 for the PEGASUS model and S1fpintsoil specification
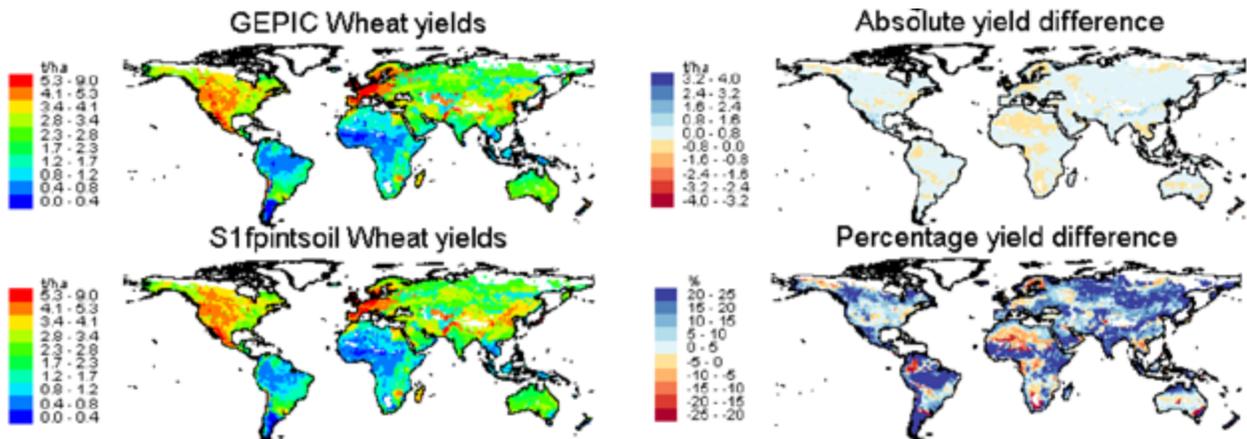
Figure G13. Irrigated wheat yields averaged over 2090–2099 for the GEPIC model and S1fpintsoil specification
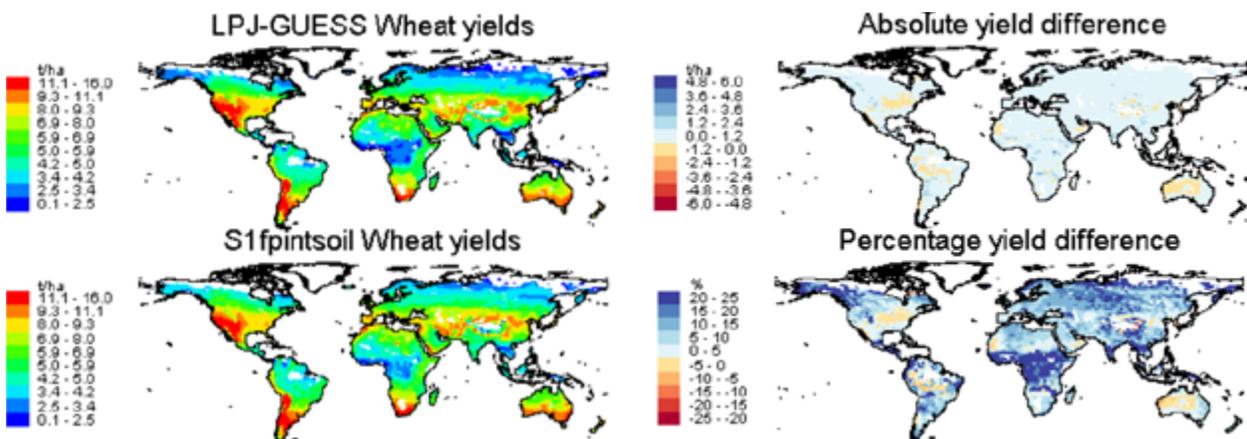
Figure G14. Irrigated wheat yields averaged over 2090–2099 for the LPJ-GUESS model and S1fpintsoil specification
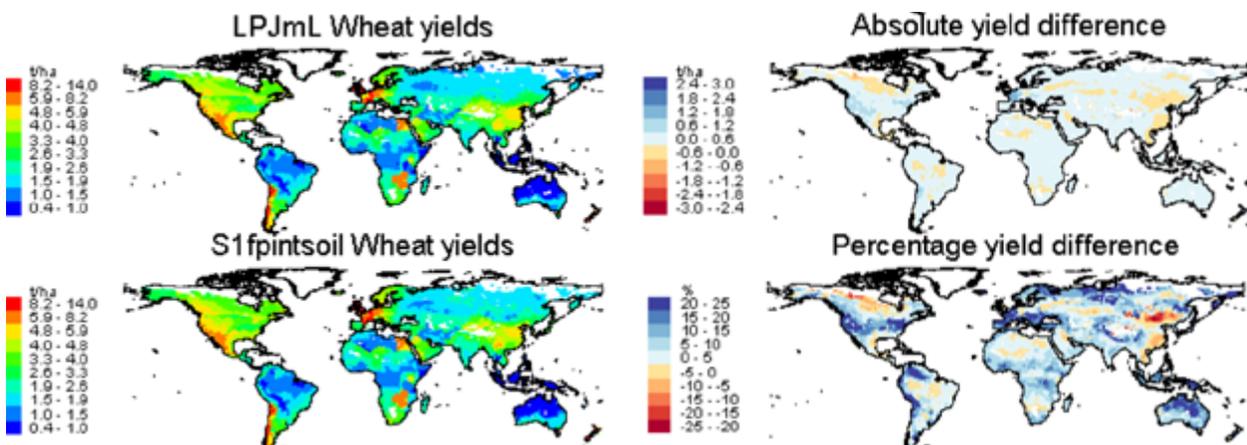
Figure G15. Irrigated wheat yields averaged over 2090–2099 for the LPJmL model and S1fpintsoil specification
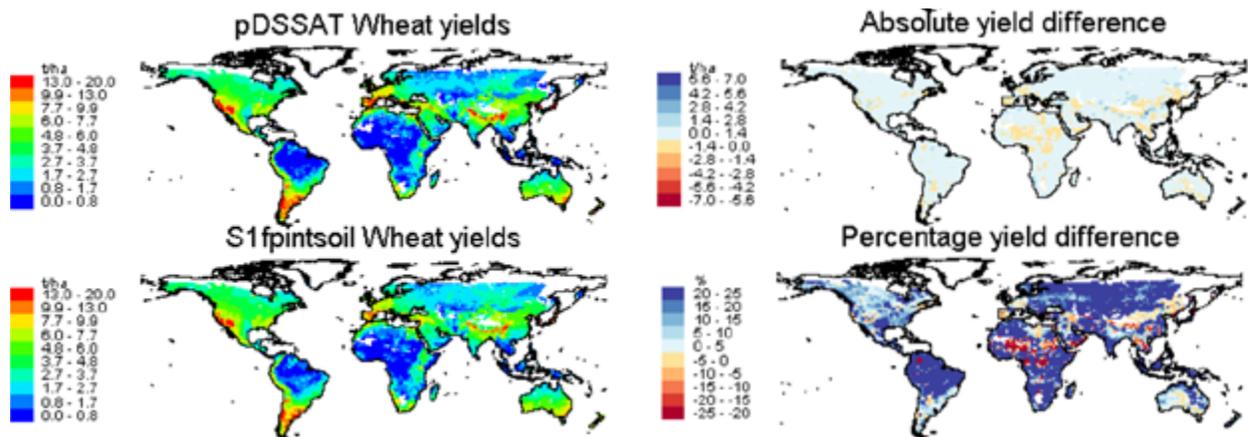
Figure G16. Irrigated wheat yields averaged over 2090–2099 for the pDSSAT model and S1fpintsoil specification
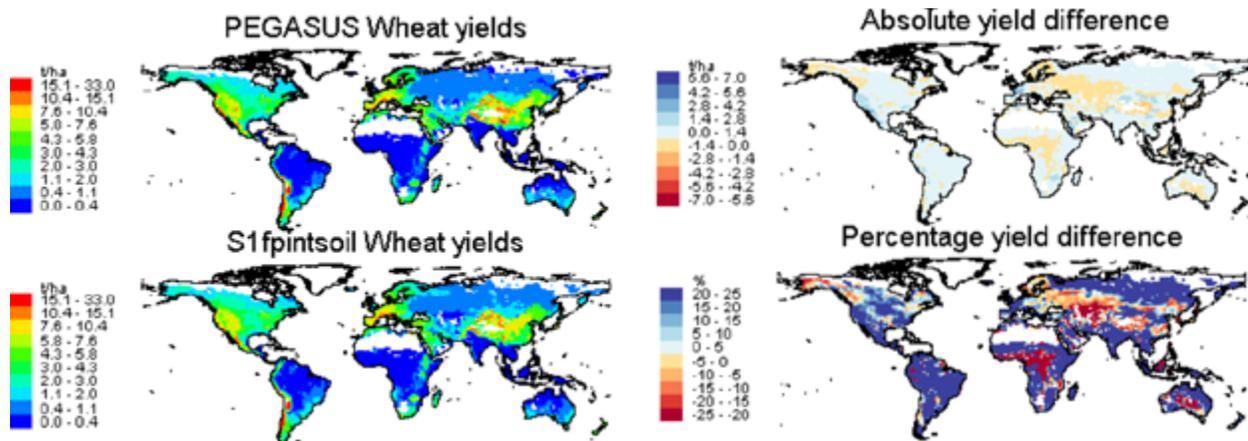
Figure G17. Irrigated wheat yields averaged over 2090–2099 for the PEGASUS model and S1fpintsoil specification

Appendix H. In-sample validation for *PIRRWW* (S1fpintsoil specification)
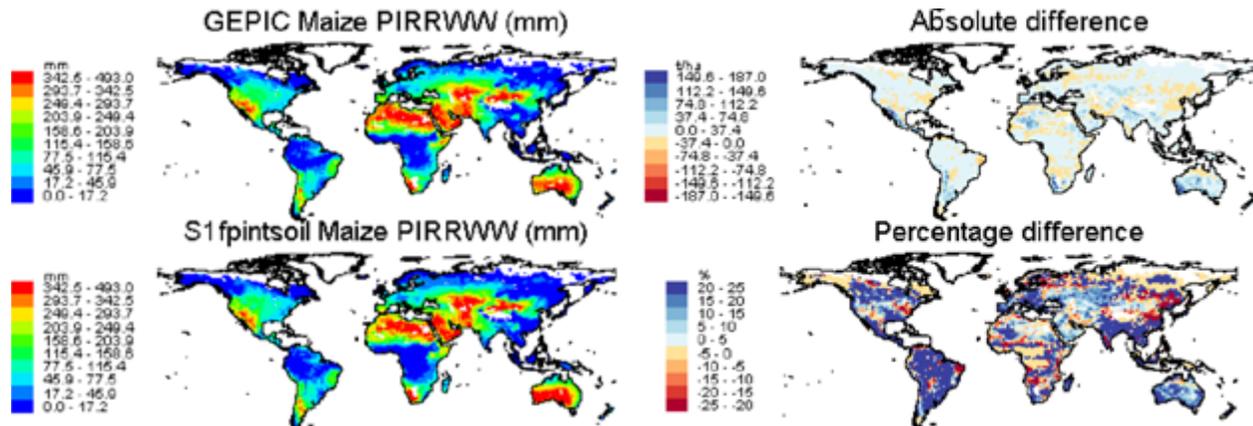
Figure H1. Irrigation water withdrawal for maize averaged over 2090–2099 for the GEPIC model and S1fpintsoil specification
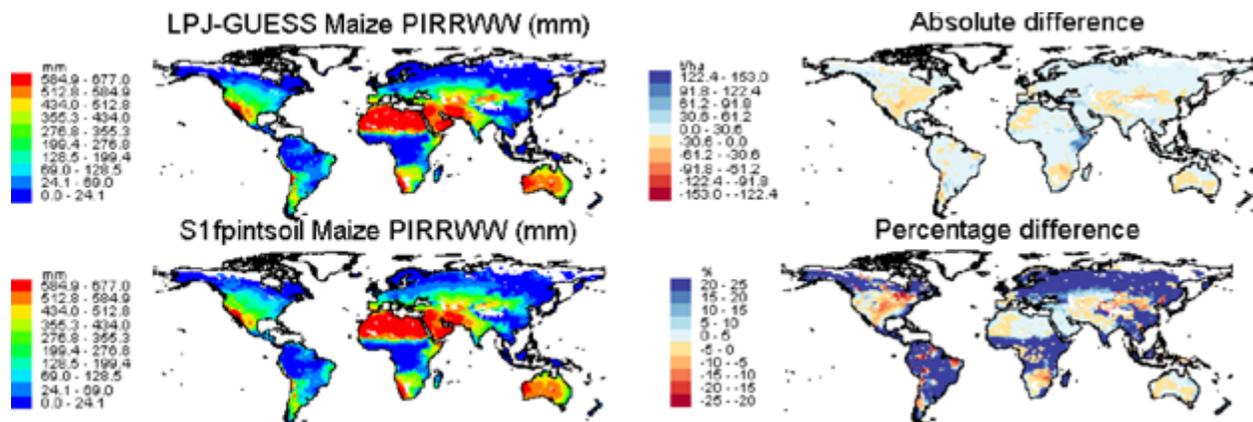
Figure H2. Irrigation water withdrawal for maize averaged over 2090–2099 for the LPJ-GUESS model and S1fpintsoil specification
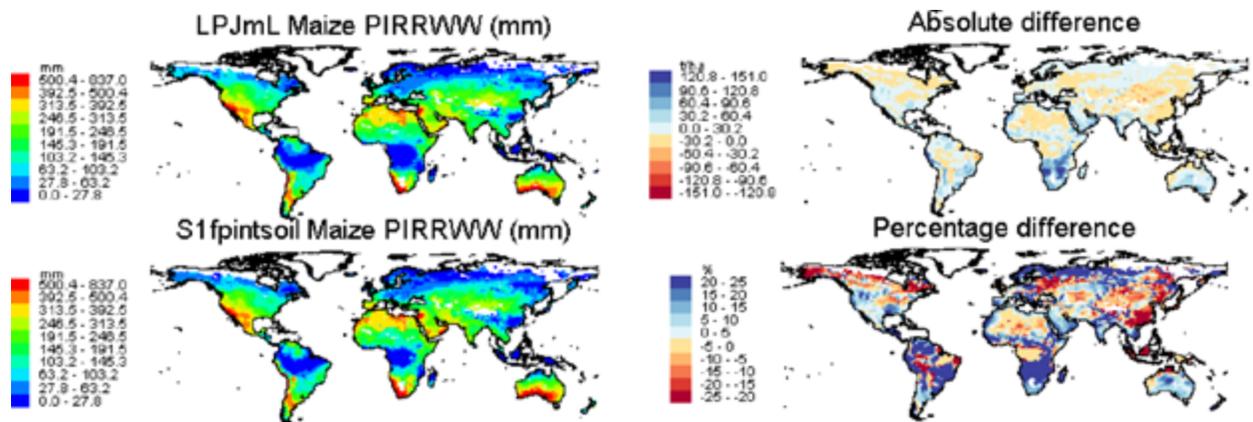
Figure H3. Irrigation water withdrawal for maize averaged over 2090–2099 for the LPJmL model and S1fpintsoil specification
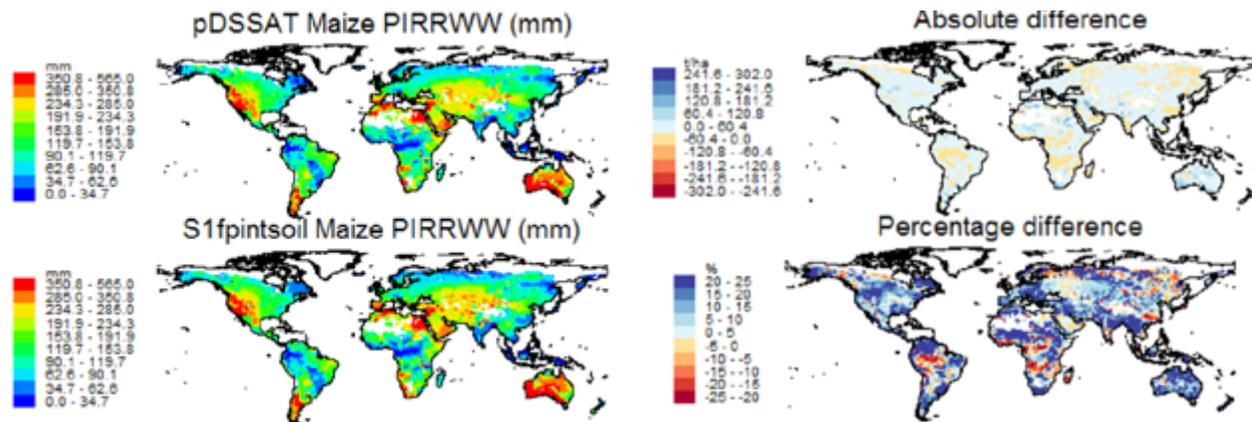
Figure H4. Irrigation water withdrawal for maize averaged over 2090–2099 for the pDSSAT model and S1fpintsoil specification
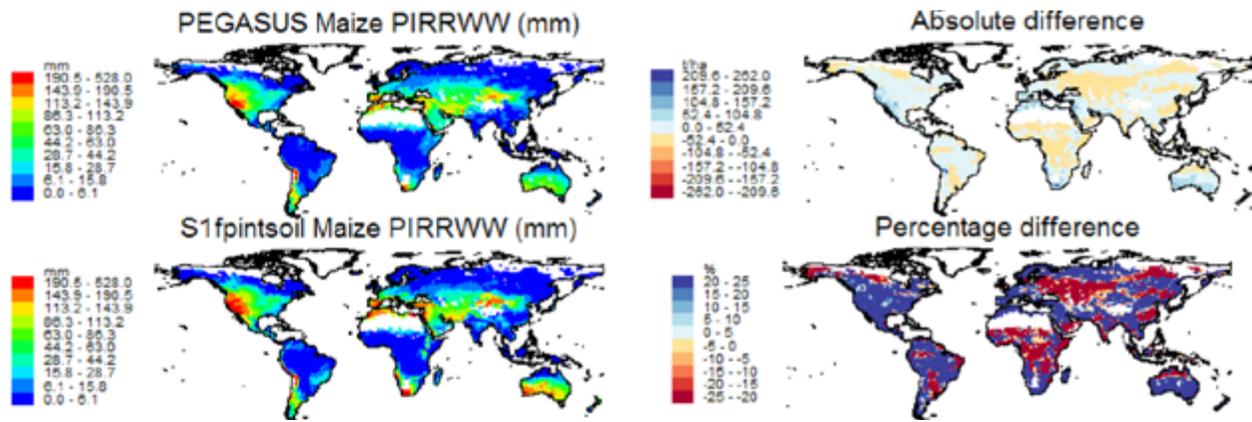
Figure H5. Irrigation water withdrawal for maize averaged over 2090–2099 for the PEGASUS model and S1fpintsoil specification
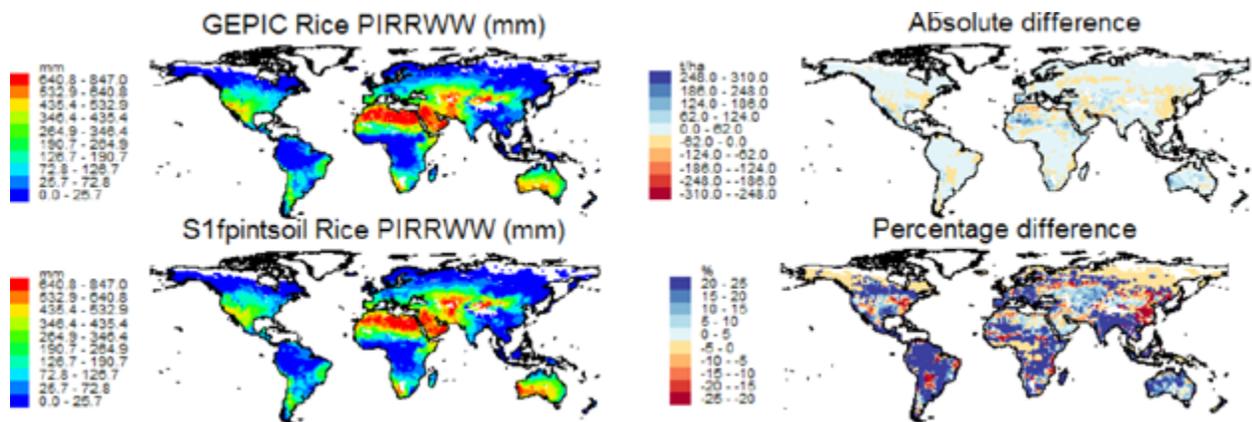
Figure H6. Irrigation water withdrawal for rice averaged over 2090–2099 for the GEPIC model and S1fpintsoil specification
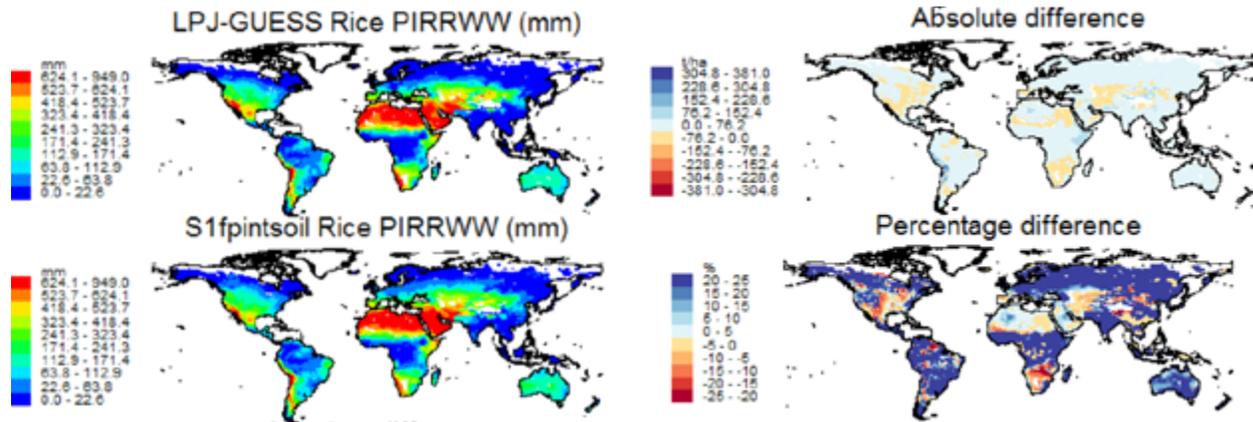
Figure H7. Irrigation water withdrawal for rice averaged over 2090–2099 for the LPJ-GUESS model and S1fpintsoil specification
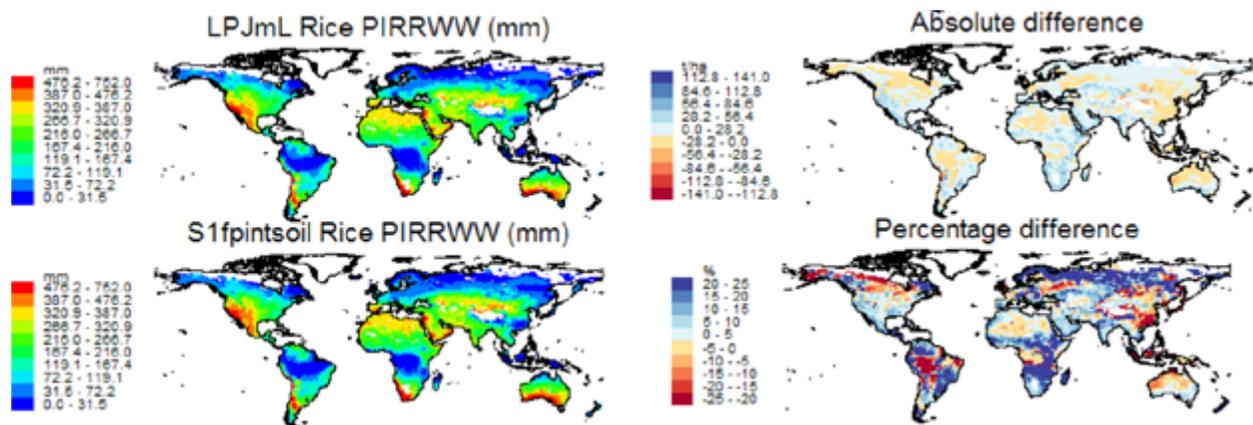
Figure H8. Irrigation water withdrawal for rice averaged over 2090–2099 for the LPJmL model and S1fpintsoil specification
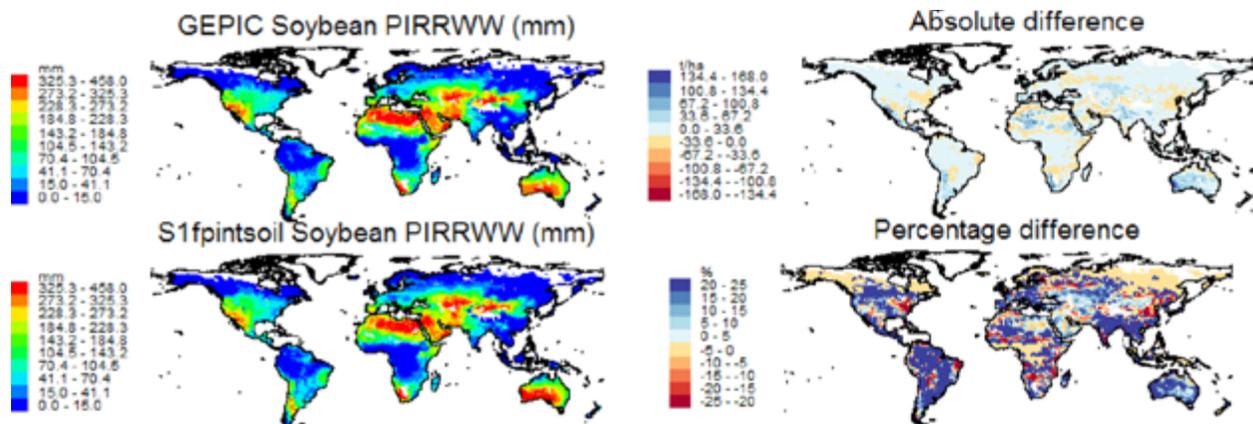
Figure H9. Irrigation water withdrawal for soybean averaged over 2090–2099 for the GEPIC model and S1fpintsoil specification
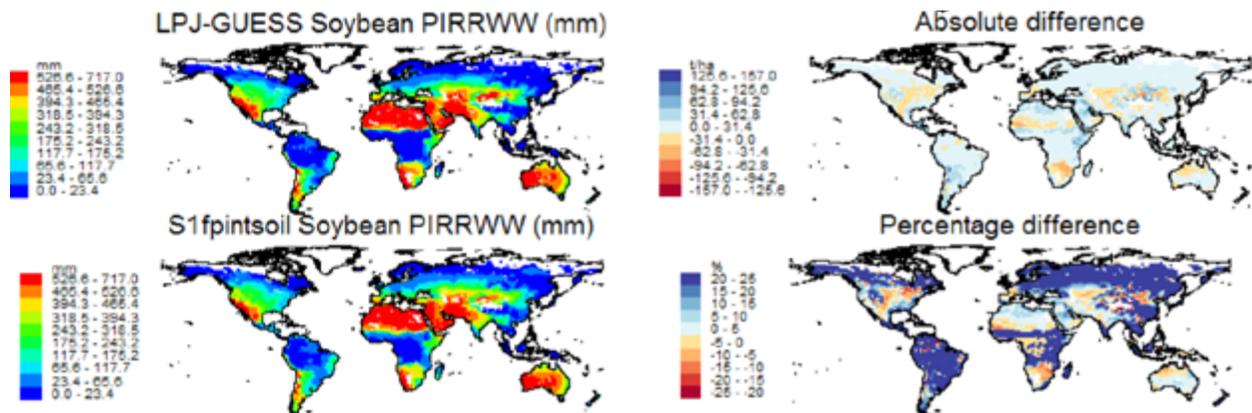
Figure H10. Irrigation water withdrawal for soybean averaged over 2090–2099 for the LPJ-GUESS model and S1fpintsoil specification
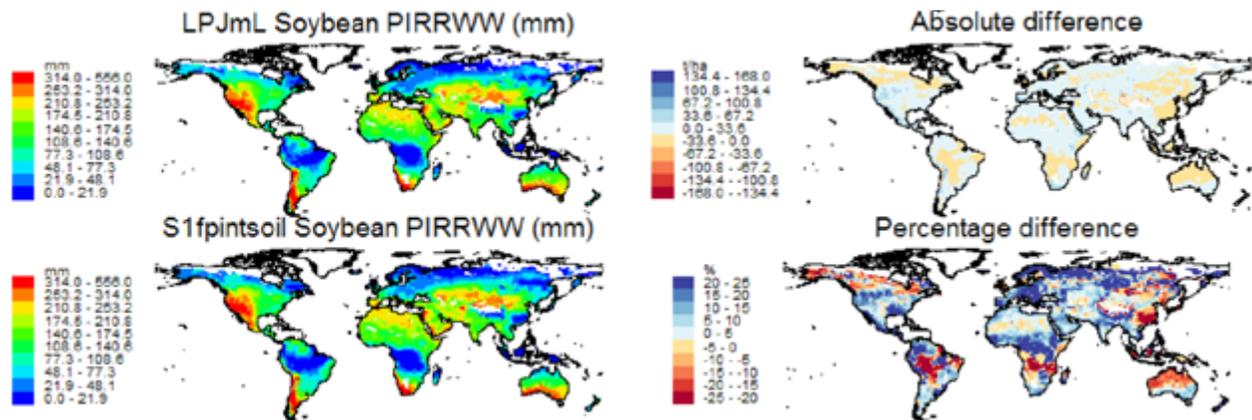
Figure H11. Irrigation water withdrawal for soybean averaged over 2090–2099 for the LPJmL model and S1fpintsoil specification
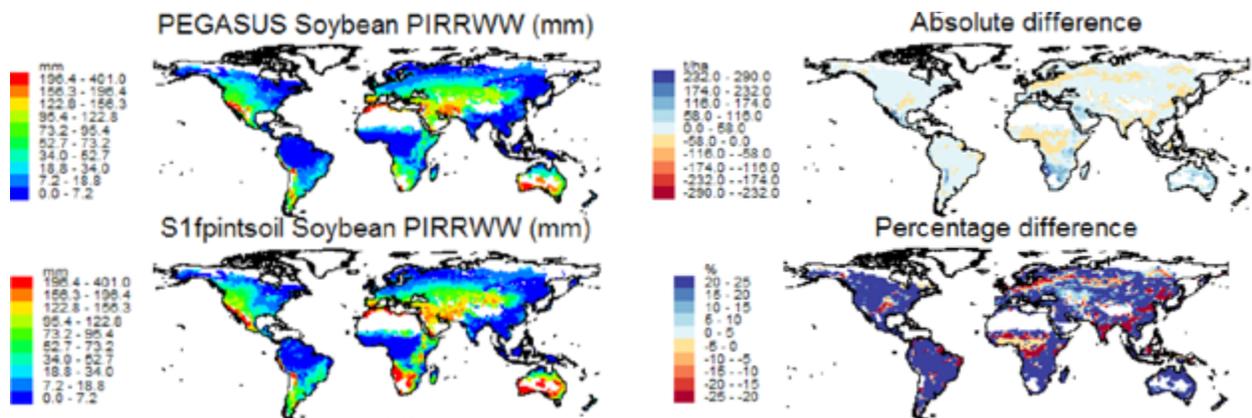
Figure H12. Irrigation water withdrawal for soybean averaged over 2090–2099 for the PEGASUS model and S1fpintsoil specification
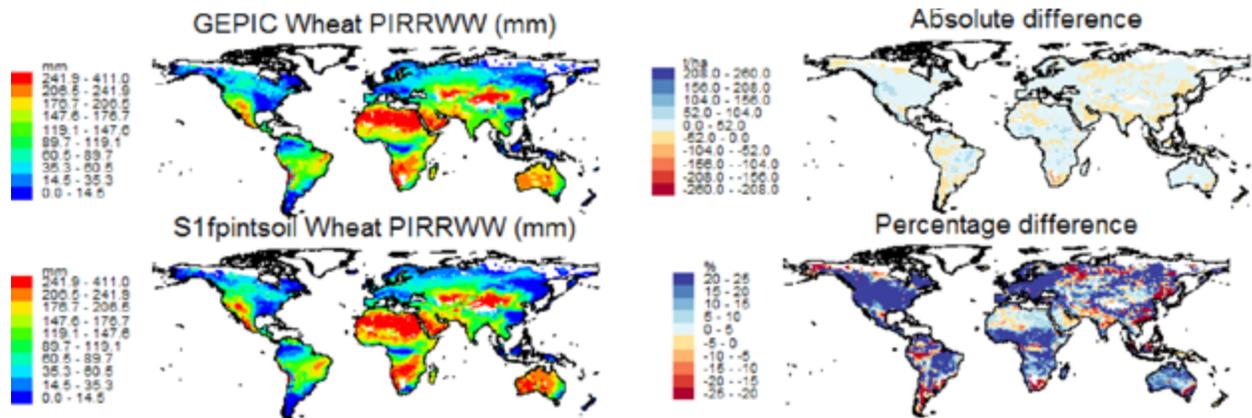
Figure H13. Irrigation water withdrawal for wheat averaged over 2090–2099 for the GEPIC model and S1fpintsoil specification
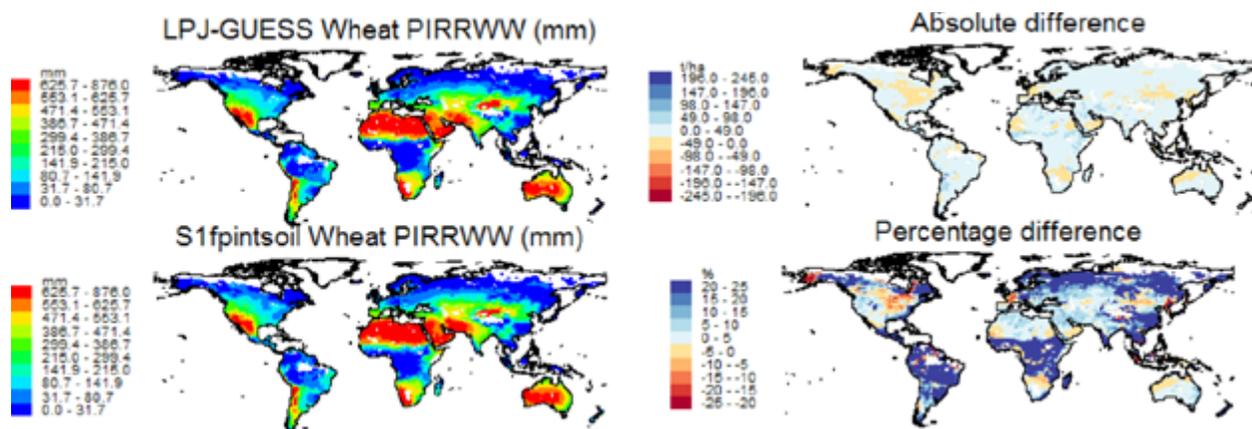
Figure H14. Irrigation water withdrawal for wheat averaged over 2090–2099 for the LPJ-GUESS model and S1fpintsoil specification
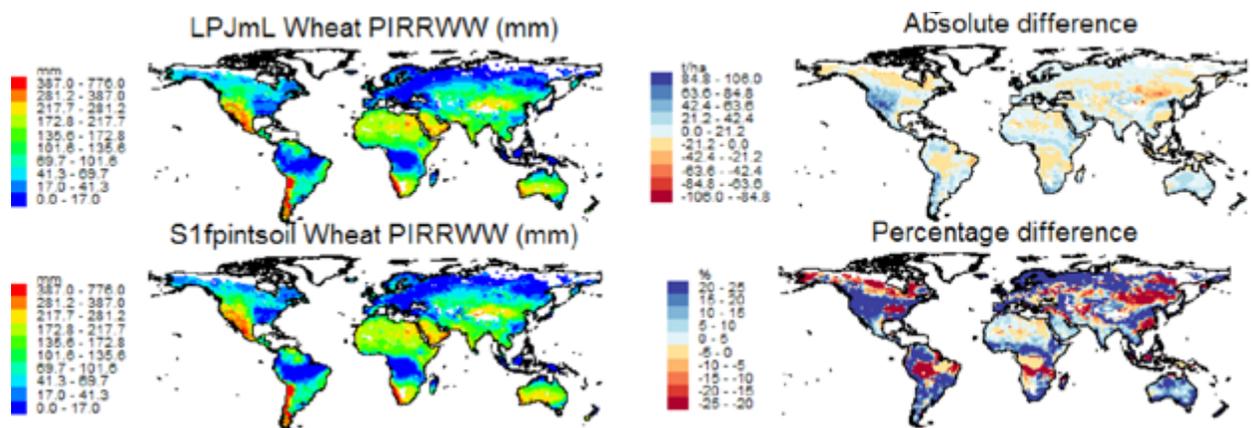
Figure H15. Irrigation water withdrawal for wheat averaged over 2090–2099 for the LPJmL model and S1fpintsoil specification
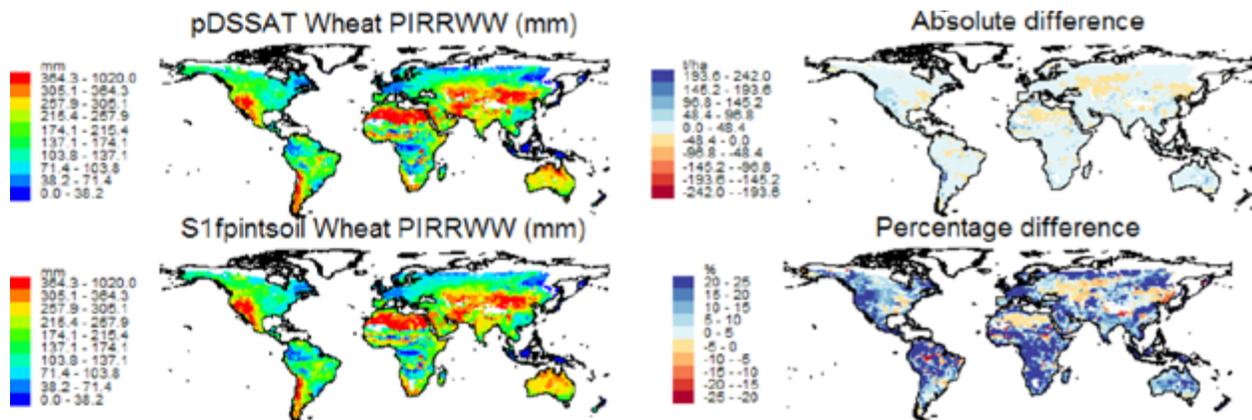
Figure H16. Irrigation water withdrawal for wheat averaged over 2090–2099 for the pDSSAT model and S1fpintsoil specification
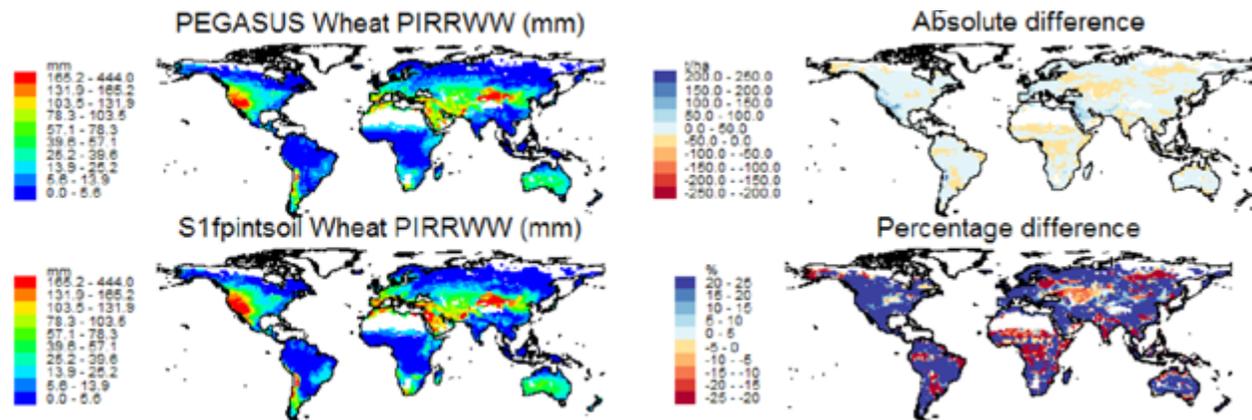
Figure H17. Irrigation water withdrawal for wheat averaged over 2090–2099 for the PEGASUS model and S1fpintsoil specification

Joint Program Report Series - Recent Articles

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