

The GGCMI Phase II experiment: global crop yield responses to changes in carbon dioxide, temperature, water, and nitrogen levels

James Franke^{a,b,*}, Joshua Elliott^{b,c}, Christoph Müller^d, Alexander Ruane^e, Abigail Snyder^f, Jonas Jägermeyr^{c,b,d,e}, Juraj Balkovic^{g,h}, Philippe Ciais^{i,j}, Marie Dury^k, Pete Falloon^l, Christian Folberth^g, Louis François^k, Tobias Hank^m, Munir Hoffmannⁿ, Cesar Izaurralde^{o,p}, Ingrid Jacquemin^k, Curtis Jones^o, Nikolay Khabarov^g, Marian Kochⁿ, Michelle Li^{b,l}, Wenfeng Liu^{r,i}, Stefan Olin^s, Meridel Phillips^{e,t}, Thomas Pugh^{u,v}, Ashwan Reddy^o, Xuhui Wang^{i,j}, Karina Williams^l, Florian Zabel^m, Elisabeth Moyer^{a,b}

^aDepartment of the Geophysical Sciences, University of Chicago, Chicago, IL, USA

^bCenter for Robust Decision-making on Climate and Energy Policy (RDCEP), University of Chicago, Chicago, IL, USA

^cDepartment of Computer Science, University of Chicago, Chicago, IL, USA

^dPotsdam Institute for Climate Impact Research, Leibniz Association (Member), Potsdam, Germany

^eNASA Goddard Institute for Space Studies, New York, NY, United States

^fJoint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD, USA

^gEcosystem Services and Management Program, International Institute for Applied Systems Analysis, Laxenburg, Austria

^hDepartment of Soil Science, Faculty of Natural Sciences, Comenius University in Bratislava, Bratislava, Slovak Republic

ⁱLaboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, 91191 Gif-sur-Yvette, France

^jSino-French Institute of Earth System Sciences, College of Urban and Environmental Sciences, Peking University, Beijing, China

^kUnité de Modélisation du Climat et des Cycles Biogéochimiques, UR SPHERES, Institut d'Astrophysique et de Géophysique, University of Liège, Belgium

^lMet Office Hadley Centre, Exeter, United Kingdom

^mDepartment of Geography, Ludwig-Maximilians-Universität, Munich, Germany

ⁿGeorg-August-University Göttingen, Tropical Plant Production and Agricultural Systems Modelling, Göttingen, Germany

^oDepartment of Geographical Sciences, University of Maryland, College Park, MD, USA

^pTexas AgriLife Research and Extension, Texas A&M University, Temple, TX, USA

^qDepartment of Statistics, University of Chicago, Chicago, IL, USA

^rEAWAG, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

^sDepartment of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden

^tEarth Institute Center for Climate Systems Research, Columbia University, New York, NY, USA

^uKarlsruhe Institute of Technology, IMK-IFU, 82467 Garmisch-Partenkirchen, Germany.

^vSchool of Geography, Earth and Environmental Science, University of Birmingham, Birmingham, UK.

Abstract

Concerns about food security under climate change have motivated efforts to better understand the future changes in yields by using detailed process-based models in agronomic sciences. Results of these models are often used to inform subsequent model-based analyses in agricultural economics and integrated assessment, but the computational requirements of process-based models sometimes hamper the ability to utilize these models in existing impact assessment frameworks at high resolutions globally. Conducting a large simulation set with perturbations in atmospheric CO₂ concentrations, temperature, precipitation, and nitrogen inputs, Phase II of the Global Gridded Crop Model Intercomparison (GGCMI), an activity of the Agricultural Model Intercomparison and Improvement Project (AgMIP), constitutes an unprecedentedly data-rich basis of projected yield changes across 12 models and five crops (maize, soy, rice, spring wheat, and winter wheat) using global gridded simulations. Results from Phase II of the GGCMI effort, a targeted experiment aimed at understanding the interaction between multiple climate variables (as well as management) are presented first. Then the construction of an “emulator or statistical representation of the simulated 30-year mean climatological output in each location is outlined. The emulator captures the response of the process-based models in a lightweight, computationally tractable form that facilitates model comparison as well as potential applications in subsequent modeling efforts such as integrated assessment.

Keywords: climate change, food security, model emulation, AgMIP, crop model

1. Introduction

2 Projecting crop yield response to a changing climate is of
3 great importance, especially as the global food production sys-
4 tem will face pressure from increased demand over the next
5 century. Climate-related reductions in supply could therefore
6 have severe socioeconomic consequences. Multiple studies
7 with different crop or climate models predict sharp reduction in
8 yields on currently cultivated cropland under business-as-usual
9 climate scenarios, although their yield projections show consid-
10 erable spread (e.g. Porter et al. (IPCC), 2014, Rosenzweig et al.,
11 2014, Schauberger et al., 2017, and references therein). Model
12 differences are unsurprising because crop responses in models
13 can be complex, with crop growth a function of complex inter-
14 actions between climate inputs and management practices.

15 Computational Models have been used to project crop yields
16 since the 1950's, beginning with statistical models (Heady,
17 1957, Heady & Dillon, 1961) that attempt to capture the rela-
18 tionship between input factors and resultant yields. These sta-
19 tistical models were typically developed on a small scale for lo-
20 cations with extensive histories of yield data. The emergence of
21 computers allowed development of numerical models that sim-
22 ulate the process of photosynthesis and the biology and phe-
23 nology of individual crops (first proposed by de Wit (1957),
24 Duncan et al. (1967) and attempted by Duncan (1972)). His-
25 torical mapping of crop model development can be found in
26 the appendix/supplementary of Rosenzweig et al. (2014). A
27 half-century of improvement in both models and computing re-
28 sources means that researchers can now run crop simulation
29 models for many years at high spatial resolution on the global
30 scale.

31 Both types of models continue to be used, and compara-
32 tive studies have concluded that when done carefully, both ap-

33 proaches can provide similar yield estimates (e.g. Lobell &
34 Burke, 2010, Moore et al., 2017, Roberts et al., 2017, Zhao
35 et al., 2017). Models tend to agree broadly in major response
36 patterns, including a reasonable representation of the spatial
37 pattern in historical yields of major crops (e.g. Elliott et al.,
38 2015, Müller et al., 2017) and projections of decreases in yield
39 under future climate scenarios.

40 Process models do continue to struggle with some important
41 details, including reproducing historical year-to-year variabil-
42 ity (e.g. Müller et al., 2017), reproducing historical yields when
43 driven by reanalysis weather (e.g. Glotter et al., 2014), and low
44 sensitivity to extreme events (e.g. Glotter et al., 2015). These
45 issues are driven in part by the diversity of new cultivars and ge-
46 netic variants, which outstrips the ability of academic modeling
47 groups to capture them (e.g. Jones et al., 2017). Models do not
48 simulate many additional factors affecting production, includ-
49 ing pests/diseases/weeds. For these reasons, individual stud-
50 ies must generally re-calibrate models to ensure that short-term
51 predictions reflect current cultivar mixes, and long-term pro-
52 jections retain considerable uncertainty (Wolf & Oijen, 2002,
53 Jagtap & Jones, 2002, Iizumi et al., 2010, Angulo et al., 2013,
54 Asseng et al., 2013, 2015). Inter-model discrepancies can also
55 be high in areas not yet cultivated (e.g. Challinor et al., 2014,
56 White et al., 2011). Finally, process-based models present ad-
57 dditional difficulties for high-resolution global studies because
58 of their complexity and computational requirements. For eco-
59 nomic impacts assessments, it is often impossible to integrate a
60 set of process-based crop models directly into an integrated as-
61 sessment model to estimate the potential cost of climate change
62 to the agricultural sector.

63 Nevertheless, process-based models are necessary for under-
64 standing the global future yield impacts of climate change for
65 many reasons. First, cultivation may shift to new areas, where
66 no yield data are currently available and therefore statistical

*Corresponding author at: 4734 S Ellis, Chicago, IL 60637, United States.
email: jfranke@uchicago.edu

models cannot apply. Yield data are also often limited in the developing world, where future climate impacts may be the most critical. Second, only process-based models can capture the growth response to elevated CO₂, novel conditions that are not represented in historical data (e.g. Pugh et al., 2016, Roberts et al., 2017). Similarly process-based models can represent novel changes in management practices (e.g. fertilizer input) that may ameliorate climate-induced damages.

Statistical emulation of crop simulations has been used to combine advantageous features of both statistical and process-based models. The statistical representation of complicated numerical simulation (e.g. O'Hagan, 2006, Conti et al., 2009), in which simulation output acts as the training data for a statistical model, has been of increasing interest with the growth of simulation complexity and volume of output. Such emulators or "surrogate models" have been used in a variety of fields including hydrology (e.g. Razavi et al., 2012), engineering (e.g. Storlie et al., 2009), environmental sciences (e.g. Ratto et al., 2012), and climate (e.g. Castruccio et al., 2014, Holden et al., 2014). For agricultural impacts studies, emulation of process-based models allows exploring crop yields in regions outside ranges of current cultivation and with input variables outside historical precedents, in a lightweight, flexible form that is compatible with economic studies.

In the past decade, many studies have developed emulators of crop yields from process-based models. Early studies proposing or describing potential emulators include Howden & Crimp (2005), Räisänen & Ruokolainen (2006) and Lobell & Burke (2010). In an early application, Ferrise et al. (2011) used a Artificial Neural Net trained on simulation outputs to predict wheat yields in the Mediterranean. Studies developing single-model emulators include Holzkämper et al. (2012) for the CropSyst model, Ruane et al. (2013) for the CERES wheat model, Oyebamiji et al. (2015) for the LPJmL model (for multiple crops,

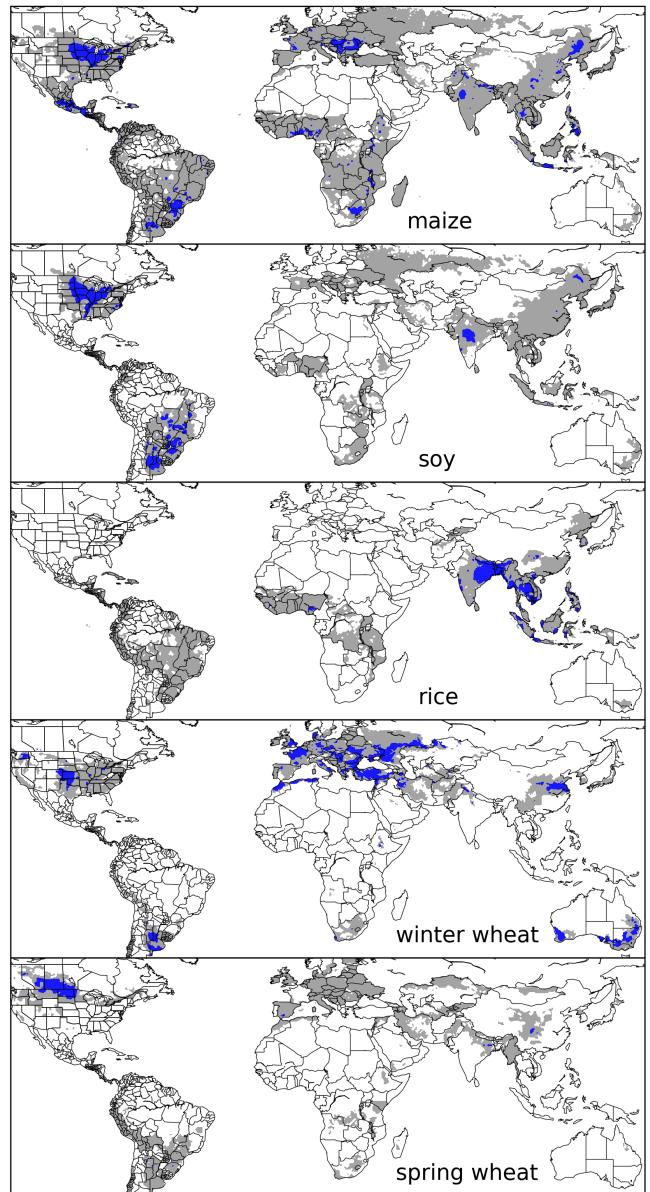


Figure 1: Presently cultivated area for rain-fed crops in the real world. Blue contour area indicates grid cells with more than 20,000 hectares (approx. 10% of the land in a grid cell near the equator for example) of crop cultivated. Gray contour shows area with more than 10 hectares cultivated. Cultivated areas for maize, rice, and soy are taken from the MIRCA2000 ("monthly irrigated and rain-fed crop areas around the year 2000") dataset (Portmann et al., 2010). Areas for winter and spring wheat areas are adapted from MIRCA2000 data and sorted by growing season. For analogous figure of irrigated crops, see Figure S1.

using multiple scenarios as a training set). In recent years, emulators have begun to be used in the context of multi-model intercomparisons, with Blanc & Sultan (2015), Blanc (2017), Ostberg et al. (2018) and Mistry et al. (2017) using them to analyze the 5 crop models of the Inter-Sectoral Impacts Model Intercomparison Project (ISIMIP) (Warszawski et al., 2014) (for

maize, soy, wheat, and rice). Approaches differ: Blanc & Sultan¹³⁵ (2015) and Blanc (2017) used local weather variables (and CO₂¹³⁶ values) and yields but emulate across soil types using historical¹³⁷ simulations and a future climate scenario (RCP8.5 over multiple¹³⁸ climate models); Ostberg et al. (2018) used global mean¹³⁹ temperature change (and CO₂) as regressors but pattern-scale¹⁴⁰ to emulate local yields using multiple climate scenarios; Mistry¹⁴¹ et al. (2017) used local weather and yields and a historical simulation¹⁴² and compare with data. Other studies have used the development¹⁴³ of emulators (or response surface) to analyze non-RCP¹⁴⁴ crop model simulations that sampled a suite of climate¹⁴⁵ (and management) perturbations: Makowski et al. (2015) for temperature,¹⁴⁶ CO₂, and nitrogen, Pirttioja et al. (2015) and Snyder et al.¹⁴⁷ (2018) for temperature, water, and CO₂, and (Fronzek¹⁴⁸ et al., 2018) for temperature and water, with all studies simulating¹⁴⁹ selected sites for a limited number of crops.¹⁵⁰

GGCMI protocol builds on the AgMIP Coordinated Climate-Crop Modeling Project (C3MP) (Ruane et al., 2014, McDermid et al., 2015) and will contribute to the AgMIP Coordinated Global and Regional Assessments (CGRA) (Ruane et al., 2018, Rosenzweig et al., 2018).

GGCMI Phase II is designed to allow addressing goals such as understanding where highest-yield regions may shift under climate change; exploring future adaptive management strategies; understanding how interacting parameters affect crop yield; quantifying uncertainties across models and major drivers; and testing strategies for producing lightweight emulators of process-based models. In this paper, we describe the GGCMI Phase II experiments, summarize output and present initial results, demonstrate that it is tractable to emulation, and present a simple climatological emulator as a potential tool for impacts assessments.¹⁵⁰

The Global Gridded Crop Model Intercomparison (GGCMI)

Phase II experiment seeks to provide a comprehensive global dataset to allow systematically exploring how process-based crop models for the major crop respond to the main climate and management drivers and their interactions. The experiment involves running a suite of process-based crop models across historical conditions perturbed by a set of defined input parameters, and was conducted as part of the Agricultural Model Intercomparison and Improvement Project (AgMIP) (Rosenzweig et al., 2013, 2014), an international effort conducted under a framework similar to the Climate Model Intercomparison Project (CMIP) (Taylor et al., 2012, Eyring et al., 2016). The

2. Materials and Methods

2.1. GGCMI Phase II: experiment design

GGCMI Phase II is the continuation of a multi-model comparison exercise begun in 2014. The initial Phase I compared harmonized yields of 21 models for 19 crops over a historical (1980-2010) scenario with a primary goal of model evaluation (Elliott et al., 2015, Müller et al., 2017). Phase II compares simulations of 12 models for 5 crops (maize, rice, soybean, spring wheat, and winter wheat) over hundreds of scenarios in which individual climate or management inputs are adjusted from

| Input variable | Abbr. | Tested range | Unit |
|------------------|-------|--|---------------------|
| CO ₂ | C | 360, 510, 660, 810 | ppm |
| Temperature | T | -1, 0, 1, 2, 3, 4, 5*, 6 | °C |
| Precipitation | W | -50, -30, -20, -10, 0, 10, 20, 30, (and W _{inf}) | % |
| Applied nitrogen | N | 10, 60, 200 | kg ha ⁻¹ |

Table 1: Phase II input variable test levels. Temperature and precipitation values indicate the perturbations from the historical, climatology. * Only simulated by one model. W-percentage does not apply to the irrigated (W_{inf}) simulations.

161 their historical values. The reduced set of crops includes the¹⁹⁴
162 three major global cereals and the major legume and accounts¹⁹⁵
163 for over 50% of human calories (in 2016, nearly 3.5 billion tons¹⁹⁶
164 or 32% of total global crop production by weight (Food and¹⁹⁷
165 Agriculture Organization of the United Nations, 2018).

166 The major goals of GGCMI Phase II are to:

- Enhance understanding of how models work by characterizing their sensitivity to input climate and nitrogen drivers.¹⁹⁹
- Study the interactions between climate variables and nitrogen inputs in driving modeled yield impacts.²⁰⁰
- Explore differences in crop response to warming across the Earth's climate regions.²⁰¹
- Provide a dataset that allows statistical emulation of crop model responses for downstream modelers.²⁰²
- Illustrate differences in potential adaptation via growing season changes.²⁰³

177 The guiding scientific rationale of GGCMI Phase II is to pro-

178 vide a comprehensive, systematic evaluation of the response²¹²
179 of process-based crop models to different values for carbon²¹³
180 dioxide, temperature, water, and applied nitrogen (collectively²¹⁴
181 known as "CTWN"). Phase II of the GGCMI project consists²¹⁵
182 of a series of simulations, each with one or more of the CTWN²¹⁶
183 dimensions perturbed over the 31-year historical time series²¹⁷
184 (1980-2010) used in Phase I. In most cases, historical daily cli-²¹⁸
185 mate inputs are taken from the 0.5 degree NASA AgMERRA²¹⁹
186 daily gridded re-analysis product specifically designed for agri-²²⁰
187 cultural modeling, with satellite-corrected precipitation (Ruane²²¹
188 et al., 2015). Two models require sub-daily input data and use²²²
189 alternative sources. See Elliott et al. (2015) for additional de-²²³
190 tails.

191 The experimental protocol consists of 9 levels for precipita-²²⁵
192 tion perturbations, 7 for temperature, 4 for CO₂, and 3 for ap-²²⁶
193 plied nitrogen, for a total of 672 simulations for rain-fed agri-²²⁷

culture and an additional 84 for irrigated (Table 1). For irrigated simulations, soil water is held at either field capacity or, for those models that include water-log damage, at maximum beneficial level. Temperature perturbations are applied as absolute offsets from the daily mean, minimum, and maximum temperature time series for each grid cell used as inputs. Precipitation perturbations are applied as fractional changes at the grid cell level, and carbon dioxide and nitrogen levels are specified as discrete values applied uniformly over all grid cells. Note that CO₂ changes are applied independently of changes in climate variables, so that higher CO₂ is not associated with higher temperatures. An additional, identical set of scenarios (at the same C, T, W, and N levels) simulate adaptive agronomy under climate change by varying the growing season for crop production. (These adaptation simulations are not shown or analyzed here.) The resulting GGCMI data set captures a distribution of crop responses over the potential space of future climate conditions.

The 12 models included in GGCMI Phase II are all mechanistic process-based crop models that are widely used in impacts assessments (Table 2). Although some of the models shares a common base (e.g. LPJmL and LPJ-GUESS and the EPIC models), they have developed independently from this shared base, for more details on the genealogy of the models see Figure S1 in Rosenzweig et al. (2014). Differences in model structure does mean that several key factors are not standardized across the experiment, including secondary soil nutrients, carry over effects across growing years including residue management and soil moisture, and extent of simulated area for different crops. Growing seasons are identical across models, but vary by crop and by location on the globe. All stresses except factors related to nitrogen, temperature, and water (e.g. Alkalinity, salinity) are disabled. No additional nitrogen inputs, such as atmospheric deposition, are considered, but some mod-

| Model (Key Citations) | Maize | Soy | Rice | Winter Wheat | Spring Wheat | N Dim. | Simulations per Crop |
|--|-------|-----|------|--------------|--------------|--------|----------------------|
| APSIM-UGOE , Keating et al. (2003), Holzworth et al. (2014) | X | X | X | – | X | Yes | 37 |
| CARAIB , Dury et al. (2011), Pirttioja et al. (2015) | X | X | X | X | X | No | 224 |
| EPIC-IIASA , Balkovi et al. (2014) | X | X | X | X | X | Yes | 35 |
| EPIC-TAMU , Izaurrealde et al. (2006) | X | X | X | X | X | Yes | 672 |
| JULES* , Osborne et al. (2015), Williams & Falloon (2015), Williams et al. (2017) | X | X | X | – | X | No | 224 |
| GEPIC , Liu et al. (2007), Folberth et al. (2012) | X | X | X | X | X | Yes | 384 |
| LPJ-GUESS , Lindeskog et al. (2013), Olin et al. (2015) | X | – | – | X | X | Yes | 672 |
| LPJmL , von Bloh et al. (2018) | X | X | X | X | X | Yes | 672 |
| ORCHIDEE-crop , Valade et al. (2014) | X | – | X | – | X | Yes | 33 |
| pDSSAT , Elliott et al. (2014), Jones et al. (2003) | X | X | X | X | X | Yes | 672 |
| PEPIC , Liu et al. (2016a,b) | X | X | X | X | X | Yes | 130 |
| PROMET*† , Mauser & Bach (2015), Hank et al. (2015), Mauser et al. (2009) | X | X | X | X | X | Yes† | 239 |
| Totals | 12 | 10 | 11 | 9 | 12 | – | 3993 (maize) |

Table 2: Models included in the GGCMI Phase II and the number of C, T, W, and N simulations performed for rain-fed crops (“Sims per Crop”), with 672 as the maximum. “N-Dim.” indicates if simulations include varying nitrogen levels; two models omit this dimension. All models provided the same set of simulations across all modeled crops, but some omitted individual crops in some cases. (For example, APSIM did not simulate winter wheat.) Irrigated simulations are provided at the level of the other covariates for each model (for an additional 84 simulations for the fully-sampled models). Geographic extent of simulation varies to some extent within a certain model for different scenarios (672 rain-fed simulations does not necessarily equal 672 climatological yields in all areas). This geographic variance only applies for areas far outside the area of currently cultivated crops. Two models (marked with *) use non-AgMERRA climate inputs. For further details on models, see Elliott et al. (2015). †PROMET provided simulations at only two nitrogen levels so is not emulated across the nitrogen dimension.

els have individual assumptions on soil organic matter that may₂₄₆ any of four initially specified levels of participation, so the num-
release additional nitrogen through mineralization. See Rosen-₂₄₇ ber of simulations varies by model, with some sampling only a
zweig et al. (2014), Elliott et al. (2015) and Müller et al. (2017)₂₄₈ part of the experiment variable space. Most modeling groups
for further details on models and underlying assumptions.₂₄₉ simulate all five crops in the protocol, but some omitted one
or more. Table 2 provides details of coverage for each model.
Each model is run at 0.5 degree spatial resolution and covers
all currently cultivated areas and much of the uncultivated land
area. Coverage extends considerably outside currently culti-
vated areas because cultivation will likely shift under climate
change. See Figure 1 for the present-day cultivated area of
rain-fed crops, and Figure S1 in the supplemental material for
irrigated crops. Some areas such as Greenland, far-northern
Canada, Siberia, Antarctica, the Gobi and Sahara deserts, and
central Australia are not simulated as they are assumed to re-
main non-arable even under an extreme climate change. Grow-
ing seasons are standardized across models with data adapted
from several sources (Sacks et al., 2010, Portmann et al., 2008,
2010).
The participating modeling groups provide simulations at₂₆₃
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
5510
5511
5512
5513
5514
5515
5516
5517
5518
5519
5520
5521
5522
5523
5524
5525
5526
5527
5528
5529
5530
5531
5532
5533
5534
5535
5536
5537
5538
5539
55310
55311
55312
55313
55314
55315
55316
55317
55318
55319
55320
55321
55322
55323
55324
55325
55326
55327
55328
55329
55330
55331
55332
55333
55334
55335
55336
55337
55338
55339
55340
55341
55342
55343
55344
55345
55346
55347
55348
55349
55350
55351
55352
55353
55354
55355
55356
55357
55358
55359
55360
55361
55362
55363
55364
55365
55366
55367
55368
55369
55370
55371
55372
55373
55374
55375
55376
55377
55378
55379
55380
55381
55382
55383
55384
55385
55386
55387
55388
55389
55390
55391
55392
55393
55394
55395
55396
55397
55398
55399
553100
553101
553102
553103
553104
553105
553106
553107
553108
553109
553110
553111
553112
553113
553114
553115
553116
553117
553118
553119
553120
553121
553122
553123
553124
553125
553126
553127
553128
553129
553130
553131
553132
553133
553134
553135
553136
553137
553138
553139
553140
553141
553142
553143
553144
553145
553146
553147
553148
553149
553150
553151
553152
553153
553154
553155
553156
553157
553158
553159
553160
553161
553162
553163
553164
553165
553166
553167
553168
553169
553170
553171
553172
553173
553174
553175
553176
553177
553178
553179
553180
553181
553182
553183
553184
553185
553186
553187
553188
553189
553190
553191
553192
553193
553194
553195
553196
553197
553198
553199
553200
553201
553202
553203
553204
553205
553206
553207
553208
553209
553210
553211
553212
553213
553214
553215
553216
553217
553218
553219
553220
553221
553222
553223
553224
553225
553226
553227
553228
553229
553230
553231
553232
553233
553234
553235
553236
553237
553238
553239
553240
553241
553242
553243
553244
553245
553246
553247
553248
553249
553250
553251
553252
553253
553254
553255
553256
553257
553258
553259
553260
553261
553262
553263
553264
553265
553266
553267
553268
553269
553270
553271
553272
553273
553274
553275
553276
553277
553278
553279
553280
553281
553282
553283
553284
553285
553286
553287
553288
553289
553290
553291
553292
553293
553294
553295
553296
553297
553298
553299
553300
553301
553302
553303
553304
553305
553306
553307
553308
553309
553310
553311
553312
553313
553314
553315
553316
553317
553318
553319
553320
553321
553322
553323
553324
553325
553326
553327
553328
553329
553330
553331
553332
553333
553334
553335
553336
553337
553338
553339
553340
553341
553342
553343
553344
553345
553346
553347
553348
553349
553350
553351
553352
553353
553354
553355
553356
553357
553358
553359
553360
553361
553362
553363
553364
553365
553366
553367
553368
553369
553370
553371
553372
553373
553374
553375
553376
553377
553378
553379
553380
553381
553382
553383
553384
553385
553386
553387
553388
553389
553390
553391
553392
553393
553394
553395
553396
553397
553398
553399
553400
553401
553402
553403
553404
553405
553406
553407
553408
553409
553410
553411
553412
553413
553414
553415
553416
553417
553418
553419
553420
553421
553422
553423
553424
553425
553426
553427
553428
553429
553430
553431
553432
553433
553434
553435
553436
553437
553438
553439
553440
553441
553442
553443
553444
553445
553446
553447
553448
553449
553450
553451
553452
553453
553454
553455
553456
553457
553458
553459
553460
553461
553462
553463
553464
553465
553466
553467
553468
553469
553470
553471
553472
553473
553474
553475
553476
553477
553478
553479
553480
553481
553482
553483
553484
553485
553486
553487
553488
553489
553490
553491
553492
553493
553494
553495
553496
553497
553498
553499
553500
553501
553502
553503
553504
553505
553506
553507
553508
553509
553510
553511
553512
553513
553514
553515
553516
553517
553518
553519
553520
553521
553522
553523
553524
553525
553526
553527
553528
553529
553530
553531
553532
553533
553534
553535
553536
553537
553538
553539
553540
553541
553542
553543
553544
553545
553546
553547
553548
553549
553550
553551
553552
553553
553554
553555
553556
553557
553558
553559
553560
553561
553562
553563
553564
553565
553566
553567
553568
553569
553570
553571
553572
553573
553574
553575
553576
553577
553578
553579
553580
553581
553582
553583
553584
553585
553586
553587
553588
553589
553590
553591
553592
553593
553594
553595
553596
553597
553598
553599
553600
553601
553602
553603
553604
553605
553606
553607
553608
553609
553610
553611
553612
553613
553614
553615
553616
553617
553618
553619
553620
553621
553622
553623
553624
553625
553626
553627
553628
553629
553630
553631
553632
553633
553634
553635
553636
553637
553638
553639
553640
553641
553642
553643
553644
553645
553646
553647
553648
553649
553650
553651
553652
553653
553654
553655
553656
553657
553658
553659
553660
553661
553662
553663
553664
553665
553666
553667
553668
553669
553670
553671
553672
553673
553674
553675
553676
553677
553678
553679
553680
553681
553682
553683
553684
553685
553686
553687
553688
553689
553690
553691
553692
553693
553694
553695
553696
553697
553698
553699
553700
553701
553702
553703
553704
553705
553706
553707
553708
553709
553710
553711
553712
553713
553714
553715
553716
553717
553718
553719
553720
553721
553722
553723
553724
553725
553726
553727
553728
553729
553730
553731
553732
553733
553734
553735
553736
553737
553738
553739
5537340
5537341
5537342
5537343
5537344
5537345
5537346
5537347
5537348
5537349
5537350
5537351
5537352
5537353
5537354
5537355
5537356
5537357
5537358
5537359
55373510
55373511
55373512
55373513
55373514
55373515
55373516
55373517
55373518
55373519
55373520
55373521
55373522
55373523
55373524
55373525
55373526
55373527
55373528
55373529
55373530
55373531
55373532
55373533
55373534
55373535
55373536
55373537
55373538
55373539
55373540
55373541
55373542
55373543
55373544
55373545
55373546
55373547
55373548
55373549
55373550
55373551
55373552
55373553
55373554
55373555
55373556
55373557
55373558
55373559
55373560
55373561
55373562
55373563
55373564
55373565
55373566
55373567
55373568
55373569
55373570
55373571
55373572
55373573
55373574
55373575
55373576
55373577
55373578
55373579
55373580
55373581
55373582
55373583
55373584
55373585
55373586
55373587
55373588
55373589
55373590
55373591
55373592
55373593
55373594
55373595
55373596
55373597
55373598
55373599
553735100
553735101
553735102
553735103
553735104
553735105
553735106
553735107
553735108
553735109
553735110
553735111
553735112
553735113
553735114
553735115
553735116
553735117
553735118
553735119
553735120
553735121
553735122
553735123
553735124
553735125
553735126
553735127
553735128
553735129
553735130
553735131
553735132
553735133
553735134
553735135
553735136
553735137
553735138
553735139
553735140
553735141
553735142
553735143
553735144
553735145
553735146
553735147
553735148
553735149
553735150
553735151
553735152
553735153
553735154
553735155
553735156
553735157
553735158
553735159
553735160
5537

264 first phase. In the case presented here however, the models
 265 are not run on the best approximation of management levels
 266 (namely nitrogen application level) by country as with phase I.
 267 As the goals of this phase of the project are focused on under-
 268 standing the sensitivity in *change* in yield to changes in input
 269 drivers –and not to simulate historical yields as accurately as
 270 possible– no direct comparison to historical yield data can be
 271 made. Additionally, even when provided with an appropriate
 272 local nitrogen level, models simulated *potential* yields that do
 273 not include reductions from pests, weeds, or diseases. Poten-
 274 tial yields represent an ideal case that is not realized in many
 275 less industrialized areas. Finally, some models are not cali-
 276 brated as they were in phase I of the project.

277 We evaluate the models here based on the response to year-
 278 to-year temperature and precipitation variability in the histori-
 279 cal record. If the models can (somewhat) faithfully represent
 280 the the historical variability in yields (which, once detrended
 281 to account for changing management levels must be driven
 282 largely by differences in weather), then the models may pro-
 283 vide some utility in understanding the impact on mean clima-
 284 tological shifts in temperature and precipitation. Specifically,
 285 we calculate a Pearson correlation coefficient between the de-
 286 trended time series of simulations and FAO data for the period
 287 1981-2009. Validating the response to CO₂ and Nitrogen appli-
 288 cations is more difficult because real world data is not available
 289 outside of small greenhouse and field level trials.

290 2.3. Climatological-mean yield emulator design

291 We construct our emulator at the 30-year climatological³⁰⁸
 292 mean level. Blanc & Sultan (2015) and Blanc (2017) have³⁰⁹
 293 shown that a emulator of a global process-based crop simula-³¹⁰
 294 tion model can be successfully developed at the yearly scale.

295 The decision to first construct a climatological-mean yield³¹²
 296 emulator is driven by the target application for this analysis³¹³
 297 tool. Many impact modelers are not focused on the changes³¹⁴

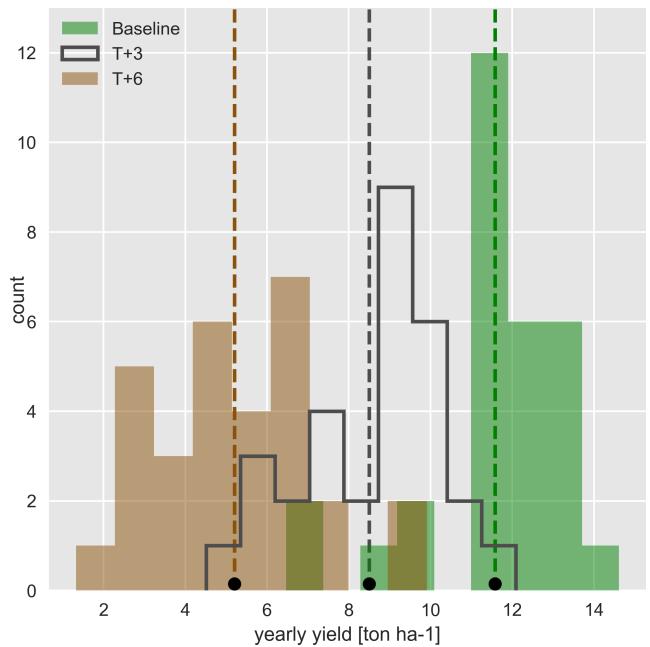


Figure 2: Example showing both climatological mean yields and distribution of yearly yields for three 30-year scenarios. Figure shows rain-fed maize for a grid cell in northern Iowa (a representative high-yield region) from the pDSSAT model, for the baseline climatology (1981-2010) and for scenarios with temperature shifted by (T) +3 and (T) +6 °C, with other variables held at baseline values. Dashed vertical lines and black dots indicate the climatological mean yield.

in the year-to-year variability in yields, but instead on the broad mean changes over the multi-decadal timescale. Emulation involves fitting individual regression models for each crop, simulation model, and 0.5 degree geographic pixel from the GGCMI Phase II data set. The regressors are the applied constant perturbations in temperature, water, nitrogen and CO₂, we aggregate the simulation outputs in the time dimension, and regress on the 30-year mean yields. (See Figure 2 for illustration). The regression therefore omits information about yield responses to year-to-year climate perturbations, which are more complex. Emulating inter-annual yield variations would likely require considering statistical details of the historical climate time series, including changes in marginal distribution and temporal dependencies. (Future work should explore this). The climatological emulation indirectly includes any yield response to geographically distributed factors such as soil type, insolation, and the baseline climate itself, because we construct separate emula-

315 tors for each grid cell. The emulator parameter matrices are
 316 portable and the yield computations are cheap even at the half-
 317 degree grid cell resolution, so we do not aggregate in space at
 318 this time.

319 Blanc & Sultan (2015) and Blanc (2017) have shown that a
 320 fractional polynomial specification is more effective than a stan-
 321 dard polynomial for representing simulations at the yearly level
 322 across different soil types geographically (not at the grid cell
 323 level). We do not test this specification here, and instead use as
 324 a starting point a standard third-order polynomial to represent
 325 the climatological-mean response at the grid cell level as it is
 326 the simplest effective specification.

327 (2009) for T and He et al. (2016) for W). We include inter-
 328 action terms (both linear and higher-order) because past stud-
 329 ies have shown them to be significant effects. For example,
 330 Lobell & Field (2007) and Tebaldi & Lobell (2008) showed
 331 that in real-world yields, the joint distribution in T and W is
 332 needed to explain observed yield variance (C and N are fixed
 333 in these data). Other observation-based studies have shown the
 334 importance of the interaction between water and nitrogen (e.g.
 335 Aulakh & Malhi, 2005), and between nitrogen and carbon diox-
 336 ide (Osaki et al., 1992, Nakamura et al., 1997). We do not focus
 337 on comparing different model specifications in this study, and
 338 instead stick to a relatively simple parameterized specification
 339 that allows for some, albeit limited, coefficient interpretation.

340 We regress climatological-mean yields against a third-order
 341 polynomial in C, T, W, and N with interaction terms. The
 342 higher-order terms are necessary to capture any nonlinear re-
 343 sponses, which are well-documented in observations for tem-
 344 perature and water perturbations (e.g. Schlenker & Roberts
 345 346 347 348

349 The limited GGCMI variable sample space means that use
 350 of the full polynomial expression described above, which has
 351 34 terms for the rain-fed case (12 for irrigated), can be prob-
 352 lematic, and can lead to over-fitting and unstable parameter es-

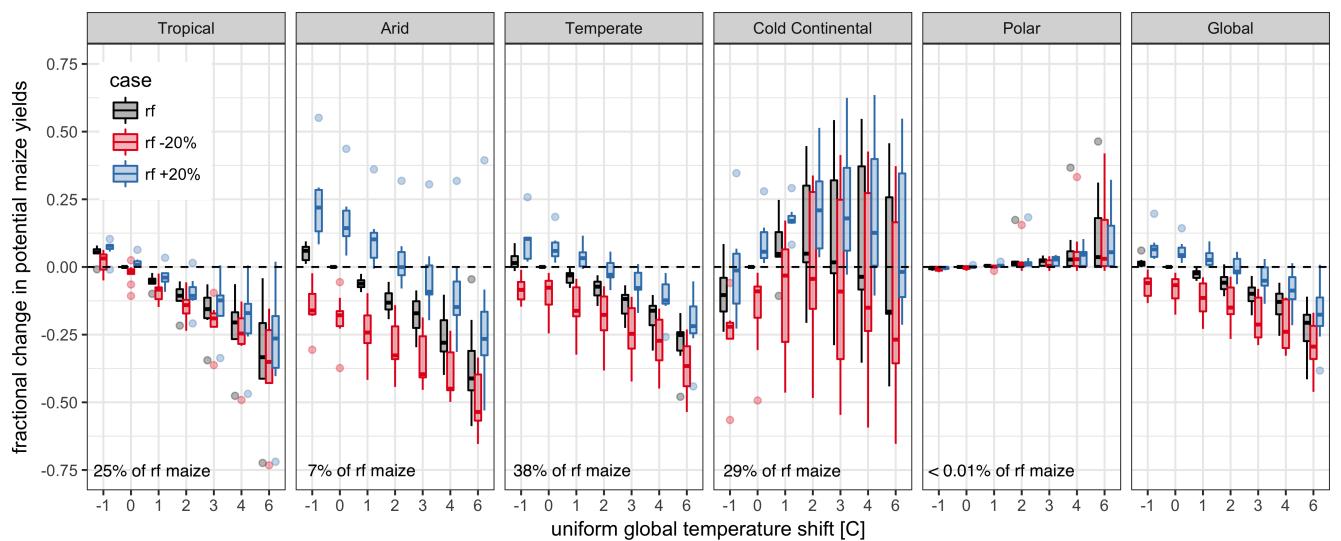


Figure 3: Illustration of the distribution of regional yield changes across the multi-model ensemble, split by Köppen-Geiger climate regions (Rubel & Kottek (2010)). We show responses of a single crop (rain-fed maize) to applied uniform temperature perturbations, for three discrete precipitation perturbation levels (rain-fed (rf) -20%, rain-fed (0), and rain-fed +20%), with CO₂ and nitrogen held constant at baseline values (360 ppm and 200 kg ha⁻¹ yr⁻¹). Y-axis is fractional change in the regional average climatological potential yield relative to the baseline. The figure shows all modeled land area; see Figure S6 in the supplemental material for only currently-cultivated land. Panel text gives the percentage of rain-fed maize presently cultivated in each climate zone (Portmann et al., 2010). Box-and-whiskers plots show distribution across models, with median marked. Box edges are first and third quartiles, i.e. box height is the interquartile range (IQR). Whiskers extend to maximum and minimum of simulations but are limited at 1.5-IQR; otherwise the outlier is shown. Models generally agree in most climate regions (other than cold continental), with projected changes larger than inter-model variance. Outliers in the tropics (strong negative impact of temperature increases) are the pDSSAT model; outliers in the high-rainfall case (strong positive impact of precipitation increases) are the JULES model. Inter-model variance increases in the case where precipitation is reduced, suggesting uncertainty in model response to water limitation. The right panel with global changes shows yield responses to an globally uniform temperature shift; note that these results are not directly comparable to simulations of more realistic climate scenarios with the same global mean change.

349 limitations. We therefore reduce the number of terms through a
 350 feature selection cross-validation process in which terms in the
 351 polynomial are tested for importance. In this procedure higher-
 352 order and interaction terms are added successively to the model;
 353 we then follow the reduction of the the aggregate mean squared
 354 error with increasing terms and eliminate those terms that do
 355 not contribute significant reductions. See supplemental docu-
 356 ments for more details. We select terms by applying the feature
 357 selection process to the three models that provided the com-
 358 plete set of 672 rain-fed simulations (pDSSAT, EPIC-TAMU,
 359 and LPJmL); the resulting choice of terms is then applied for
 360 all emulators.

361 Feature importance is remarkably consistent across all three³⁸⁴
 362 models and across all crops (see Figure S4 in the supplemental³⁸⁵
 363 material). The feature selection process results in a final poly-³⁸⁶
 364 nomial in 23 terms, with 11 terms eliminated. We omit the N^3 ³⁸⁷
 365 term, which cannot be fitted because we sample only three ni-³⁸⁸
 366 trogen levels. We eliminate many of the C terms: the cubic,³⁸⁹
 367 the CT, CTN, and CWN interaction terms, and all higher order
 368 interaction terms in C. Finally, we eliminate two 2nd-order in-³⁹⁰
 369 teraction terms in T and one in W. Implication of this choice³⁹¹
 370 include that nitrogen interactions are complex and important,³⁹²
 371 and that water interaction effects are more nonlinear than those³⁹³
 372 in temperature. The resulting statistical model (Equation 1) is³⁹⁴
 373 used for all grid cells, models, and crops:

$$\begin{aligned}
 Y = & K_1 \\
 & + K_2 C + K_3 T + K_4 W + K_5 N \\
 & + K_6 C^2 + K_7 T^2 + K_8 W^2 + K_9 N^2 \\
 & + K_{10} CW + K_{11} CN + K_{12} TW + K_{13} TN + K_{14} WN \\
 & + K_{15} T^3 + K_{16} W^3 + K_{17} TWN \\
 & + K_{18} T^2 W + K_{19} W^2 T + K_{20} W^2 N \\
 & + K_{21} N^2 C + K_{22} N^2 T + K_{23} N^2 W
 \end{aligned} \tag{1}$$

374 To fit the parameters K , we use a Bayesian Ridge probabilis-
 375 tic estimator (MacKay, 1991), which reduces volatility in pa-
 376 rameter estimates when the sampling is sparse, by weighting
 377 parameter estimates towards zero. The Bayesian Ridge method
 378 is necessary to maintain a consistent functional form across all
 379 models, and locations as the linear least squares fails to pro-
 380 vide a stable result in many cases. In the GGCMI Phase II
 381 experiment, the most problematic fits are those for models that
 382 provided a limited number of cases or for low-yield geographic
 383 regions where some modeling groups did not run all scenarios.
 384 Because we do not attempt to emulate models that provided
 385 less than 50 simulations, the lowest number of simulations em-
 386 ulated across the full parameter space is 130 (for the PEPIC
 387 model). We use the implementation of the Bayesian Ridge esti-
 388 mator from the scikit-learn package in Python (Pedregosa et al.,
 389 2011).

390 The resulting parameter matrices for all crop model emula-
 391 tors are available on request, as are the raw simulation data and
 392 a Python application to emulate yields. The yield output for a
 393 single GGCMI model that simulates all scenarios and all five
 394 crops is ~ 12.5 GB; the emulator is ~ 100 MB, a reduction by
 395 over two orders of magnitude.

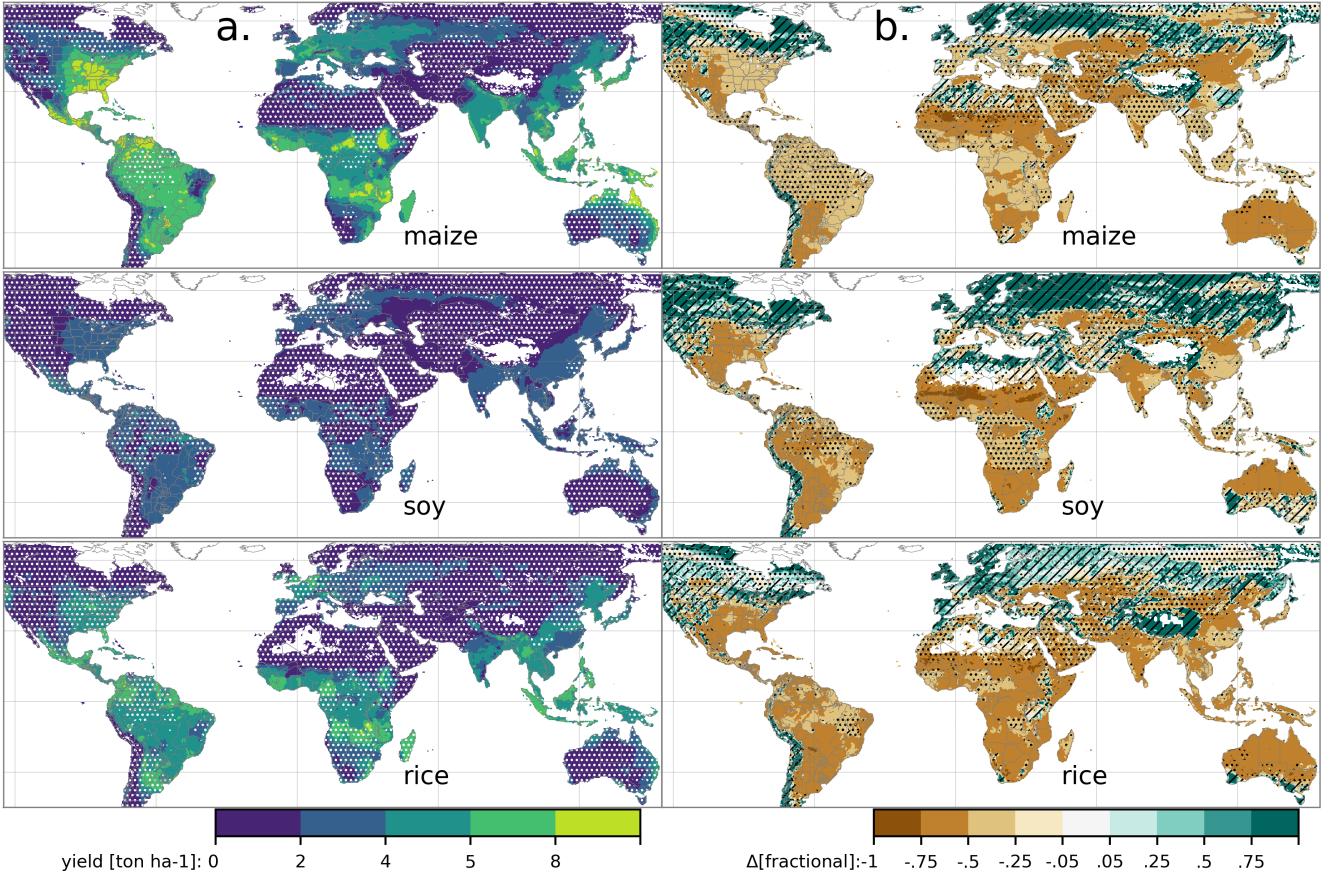


Figure 4: Illustration of the spatial pattern of potential yields and potential yield changes in the GGCMI Phase II ensemble, for three major crops. Left column (a) shows multi-model mean climatological yields for the baseline scenario for (top–bottom) rain-fed maize, soy, and rice. (For wheat see Figure S11 in the supplemental material.) White stippling indicates areas where these crops are not currently cultivated. Absence of cultivation aligns well with the lowest yield contour ($0\text{--}2 \text{ ton ha}^{-1}$). Right column (b) shows the multi-model mean fractional yield change in the extreme $T + 4^\circ\text{C}$ scenario (with other inputs at baseline values). Areas without hatching or stippling are those where confidence in projections is high: the multi-model mean fractional change exceeds two standard deviations of the ensemble. ($\Delta > 2\sigma$). Hatching indicates areas of low confidence ($\Delta < 1\sigma$), and stippling areas of medium confidence ($1\sigma < \Delta < 2\sigma$). Crop model results in cold areas, where yield impacts are on average positive, also have the highest uncertainty.

396 2.4. Emulator evaluation

397 Because no general criteria exist for defining an acceptable
 398 model emulator, we develop a metric of emulator performance
 399 specific to GGCMI. For a multi-model comparison exercise like
 400 GGCMI, a reasonable criterion is what we term the “normalized
 401 error”, which compares the fidelity of an emulator for a given
 402 model and scenario to the inter-model uncertainty. We define
 403 the normalized error e for each scenario as the difference be-
 404 tween the fractional yield change from the emulator and that in
 405 the original simulation, divided by the standard deviation of the
 406 multi-model spread (Equations 2 and 3):

$$F_{scn.} = \frac{Y_{scn.} - Y_{baseline}}{Y_{baseline}}$$

$$e_{scn.} = \frac{F_{em, scn.} - F_{sim, scn.}}{\sigma_{sim, scn.}} \quad (3)$$

Here $F_{scn.}$ is the fractional change in a model’s mean emulated or simulated yield from a defined baseline, in some scenario (scn.) in C, T, W, and N space; $Y_{scn.}$ and $Y_{baseline}$ are the absolute emulated or simulated mean yields. The normalized error e is the difference between the emulated fractional change in yield and that actually simulated, normalized by $\sigma_{sim.}$, the standard deviation in simulated fractional yields $F_{sim, scn.}$ across all models. The emulator is fit across all available simulation outputs, and then the error is calculated across the simulation scenarios provided by all nine models (Figure 8 and Figures

417 S12 and Figures S13 in supplemental documents). Note that₄₅₀
418 the normalized error e for a model depends not only on the fidelity₄₅₁
419 of its emulator in reproducing a given simulation but on₄₅₂
420 the particular suite of models considered in the intercomparison₄₅₃
421 exercise. The rationale for this choice is to relate the fidelity of₄₅₄
422 the emulation to an estimate of true uncertainty, which we take₄₅₅
423 as the multi-model spread.

424 3. Results

425 3.1. Simulation results

426 Crop models in the GGCMI ensemble show a broadly con-
427 sistent responses to climate and management perturbations in
428 most regions, with a strong negative impact of increased tem-
429 perature in all but the coldest regions. We illustrate this result
430 for rain-fed maize in Figure 3, which shows yields for the pri-
431 mary Köppen-Geiger climate regions (Rubel & Kottek, 2010).
432 In warming scenarios, models show decreases in maize yield in
433 the temperate, tropical, and arid regions that account for nearly
434 three-quarters of global maize production. These impacts are
435 robust for even moderate climate perturbations. In the temper-
436 ate zone, even a 1 degree temperature rise with other variables₄₇₀
437 held fixed leads to a median yield reduction that outweighs the
438 variance across models. A 6 degree temperature rise results in₄₇₂
439 median loss of ~25% of yields with a signal to noise of nearly₄₇₃
440 three. A notable exception is the cold continental region, where₄₇₄
441 models disagree strongly, extending even to the sign of impacts.₄₇₅
442 Model simulations of other crops produce similar responses to₄₇₆
443 warming, with robust yield losses in warmer locations and high₄₇₇
444 inter-model variance in the cold continental regions (Figures₄₇₈
445 S7).

446 The effects of rainfall changes on maize yields are also as ex-₄₈₀
447 pected and are consistent across models. Increased rainfall mit-₄₈₁
448 igates the negative effect of higher temperatures, most strongly₄₈₂
449 in arid regions. Decreased rainfall amplifies yield losses and₄₈₃

also increases inter-model variance more strongly, suggesting that models have difficulty representing crop response to water stress. We show only rain-fed maize here; see Figure S5 for the irrigated case. As expected, irrigated crops are more resilient to temperature increases in all regions, especially so where water is limiting.

Mapping the distribution of baseline yields and yield changes shows the geographic dependencies that underlie these results. Figure 4 shows baseline and changes in the T+4 scenario for rain-fed maize, soy, and rice in the multi-model ensemble mean, with locations of model agreement marked. Absolute yield potentials are have strong spatial variation, with much of the Earth's surface area unsuitable for any given crop. In general, models agree most on yield response in regions where yield potentials are currently high and therefore where crops are currently grown. Models show robust decreases in yields at low latitudes, and highly uncertain median increases at most high latitudes. For wheat crops see Figure S11; wheat projections are both more uncertain and show fewer areas of increased yield in the inter-model mean.

3.2. Simulation model validation results

Figure 5 shows the time series correlation between the simulation model yield and FAO yield data. The results are mixed, with many regions for rice and wheat being difficult to model. No single model is dominant, with each model providing near best-in-class performance in at least one location-crop combination. The presence of no vertical dark green color bars clearly illustrates the power of a multi-model intercomparison project like the one presented here. The ensemble mean yield is calculated across all 'high' nitrogen application level model simulations and correlated with the FAO data (not the mean of the correlations). The ensemble mean does not beat the best model in each case, but shows positive correlation in over 75% of the cases presented here.

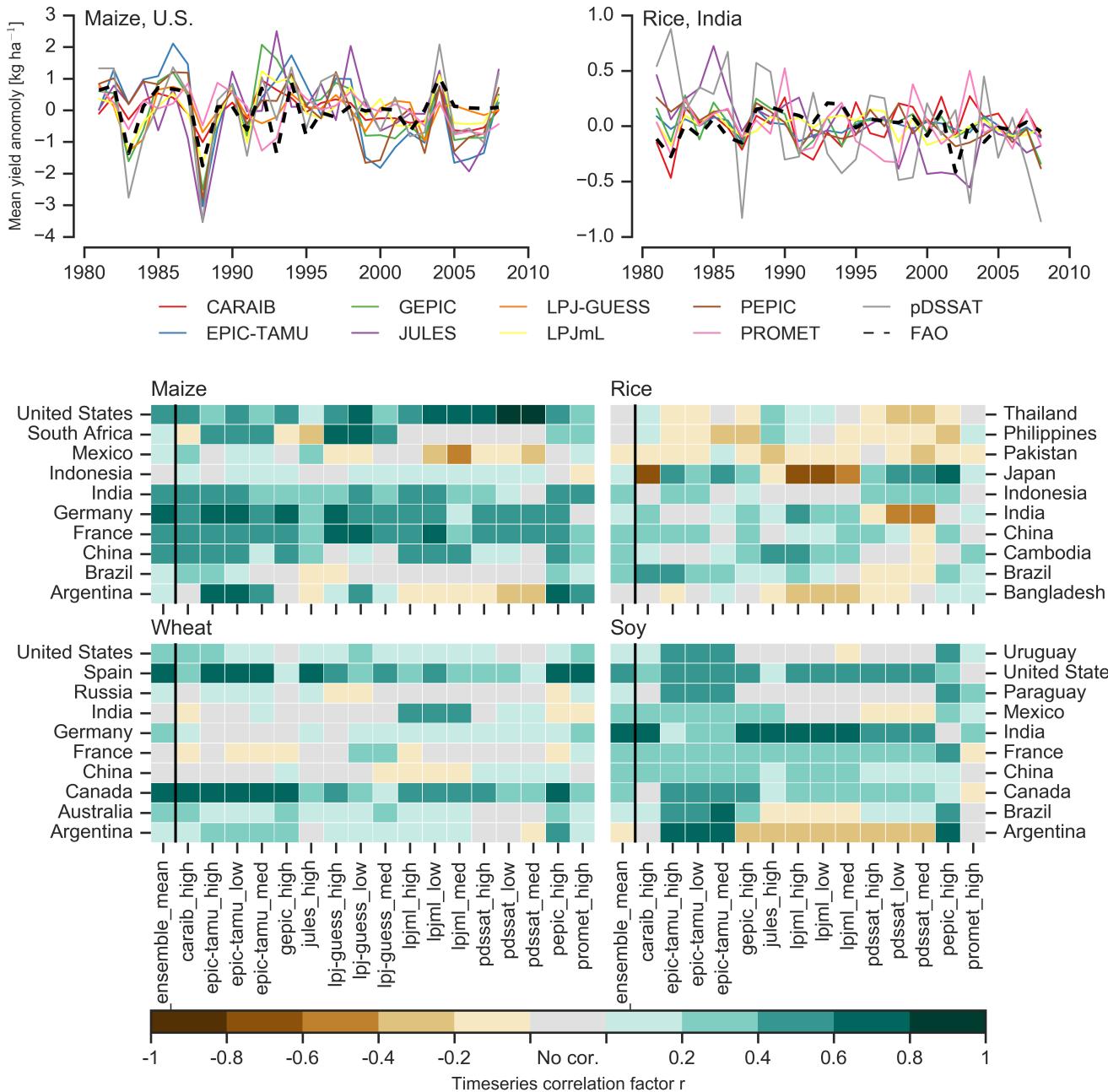


Figure 5: Time series correlation coefficients between simulated crop yield and FAO data at the country level. The top panels indicate two example cases: US maize (a good case), and rice in India (mixed case), both for the high nitrogen application case. The heatmaps illustrate the Pearson r correlation between the detrended simulation mean yield at the country level compared to the detrended FAO yield data for the top producing countries for each crop with continuous FAO data over the 1980-2010 period. Models that provided different nitrogen application levels are shown with low, med, and high label (models that did not simulate different nitrogen levels are analogous to a high nitrogen application level). The ensemble mean yield is also correlated with the FAO data.

484 Soy is qualitatively the easiest crop to represent (except in Argentina), which is likely due to the invariance of the response to nitrogen application (soy fixes atmospheric nitrogen very efficiently). Comparison to the FAO data is therefore easier than the other crops because the nitrogen application levels do not matter. US maize has the best performance across models, with nearly every model representing the historical variability to some extent. Especially good example years for US maize are 1983, 1988, and 2004 (top left panel), where every model gets the direction of the anomaly compared to surrounding years correct. 1983 and 1988 are famously bad years for US maize along with 2012 (not shown). US maize is possibly

496 both the most uniformly industrialized (in terms of management⁵³⁰
497 practices) crop and the one with the best data collection in the
498 historical period of all the cases presented here.

499 FAO data is at least one level of abstraction from ground truth⁵³¹
500 in many cases, especially in developing countries. The fail-⁵³²
501 ure of models to represent the year-to-year variability in rice in⁵³³
502 some countries in southeast Asia is likely partly due to model⁵³⁴
503 failure and partly due to lack of data. Partitioning of these con-⁵³⁵
504 tributions is impossible at this stage. Additionally, there is less⁵³⁶
505 year-to-year variability in rice yields (partially due to the frac-⁵³⁷
506 tion of irrigated cultivation). Since the Pearson r metric is scale⁵³⁸
507 invariant, it will tend to score the rice models more poorly than⁵³⁹
508 maize and soy. The pDSSAT model shows very poor perfor-⁵⁴⁰
509 mance for rice in India (top right panel).

3.3. Emulator performance

Emulation provides not only a computational tool but a means of understanding and interpreting crop yield response across the parameter space. Emulation is only possible, however, when crop yield responses are sufficiently smooth and continuous to allow fitting with a relatively simple functional form. In the GGCMI simulations, this condition largely but not always holds. Responses are quite diverse across locations, crops, and models, but in most cases local responses are regular enough to permit emulation. Figure 6 illustrates the geographic diversity of responses even in high-yield areas for a single crop and model (rain-fed maize in pDSSAT for various high-cultivation areas). This heterogeneity validates the choice of emulating at the grid cell level.

510 *One may speculate that the difference in performance be-*⁵⁴⁴
511 *tween Pakistan (no successful models) and India (many suc-*⁵⁴⁵
512 *cessful models) for rice may lie in the FAO data and not the*⁵⁴⁶
513 *models themselves. What would be so different about rice pro-*⁵⁴⁷
514 *duction across these two countries that could explain this dif-*⁵⁴⁸
515 *ference??*

516 Figure ?? shows the distribution across historical maize⁵⁵⁰
517 yields for some high producing countries. The discrepancy⁵⁵¹
518 between the simulations and FAO data is most evident in de-⁵⁵²
519 veloping nations, where nitrogen application levels are far be-⁵⁵³
520 low the 200 kg ha⁻¹ applied in the simulations shown here⁵⁵⁴
521 (though the distributions are similar in those nations other-⁵⁵⁵
522 wise). The distribution in historical yields is also calculated⁵⁵⁶
523 with the climatological-mean emulator by passing it the histor-⁵⁵⁷
524 ical (1981-2009) anomalies in growing season precipitation and⁵⁵⁸
525 temperature, CO₂ concentration of 360 ppm, and spatially vary-⁵⁵⁹
526 ing nitrogen application rates (data from: Potter et al., 2010,⁵⁶⁰
527 Mueller et al., 2012). The emulator distribution is shifted to-⁵⁶¹
528 wards the FAO distribution in cases where the nitrogen levels⁵⁶²
529 are too high in the simulations, but this does not account for⁵⁶³

Each panel in Figure 6 shows model yield output from scenarios varying only along a single dimension (CO₂, temperature, precipitation, or nitrogen addition), with other inputs held fixed at baseline levels; in all cases yields evolve smoothly across the space sampled. For reference we show the results of the full emulation fitted across the parameter space. The polynomial fit readily captures the climatological response to perturbations.

Crop yield responses generally follow similar functional forms across models, though with a spread in magnitude. Figure 7 illustrates the inter-model diversity of yield responses to the same perturbations, even for a single crop and location (rain-fed maize in northern Iowa, the same location shown in the Figure 6). The differences make it important to construct emulators separately for each individual model, and the fidelity of emulation can also differ across models. This figure illustrates a common phenomenon, that models differ more in response to perturbations in CO₂ and nitrogen perturbations than to those in temperature or precipitation. (Compare also Figures 3 and S18.) For this location and crop, CO₂ fertilization effects

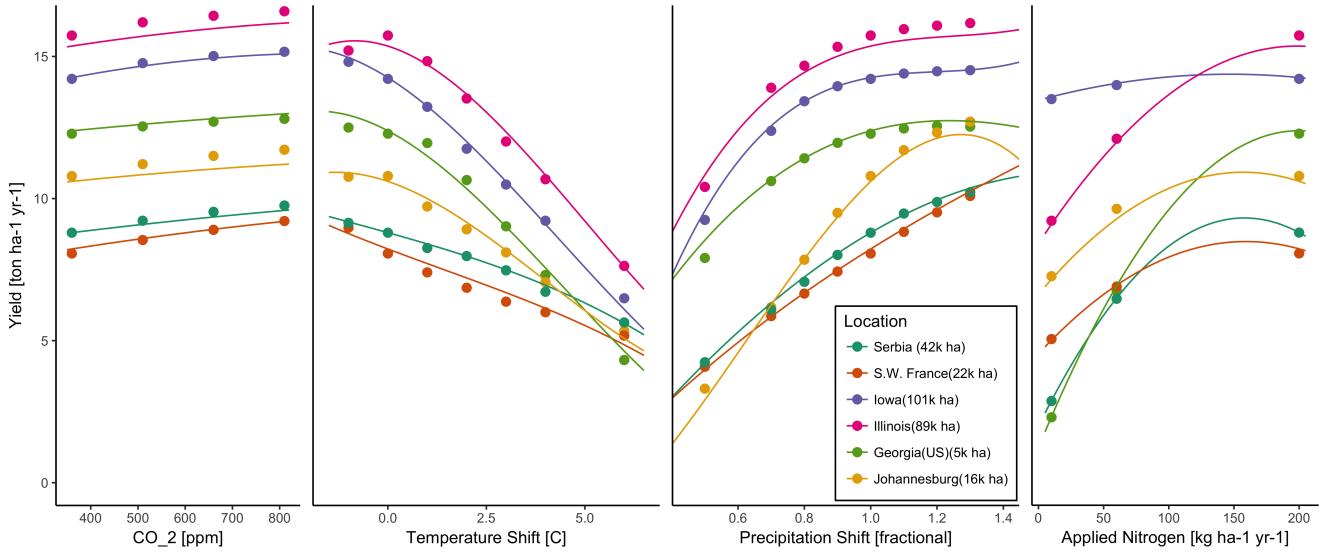


Figure 6: Example illustrating spatial variations in yield response and emulation ability. We show rain-fed maize in the pDSSAT model in six example locations selected to represent high-cultivation areas around the globe. Legend includes hectares cultivated in each selected grid cell. Each panel shows variation along a single variable, with others held at baseline values. Dots show climatological mean yield, solid lines a simple polynomial fit across this 1D variation, and dashed lines the results of the full emulator of Equation 1. The climatological response is sufficiently smooth to be represented by a simple polynomial. Because precipitation changes are imposed as multiplicative factors rather than additive offsets, locations with higher baseline precipitation have a larger absolute spread in inputs sampled than do drier locations (not visible here).

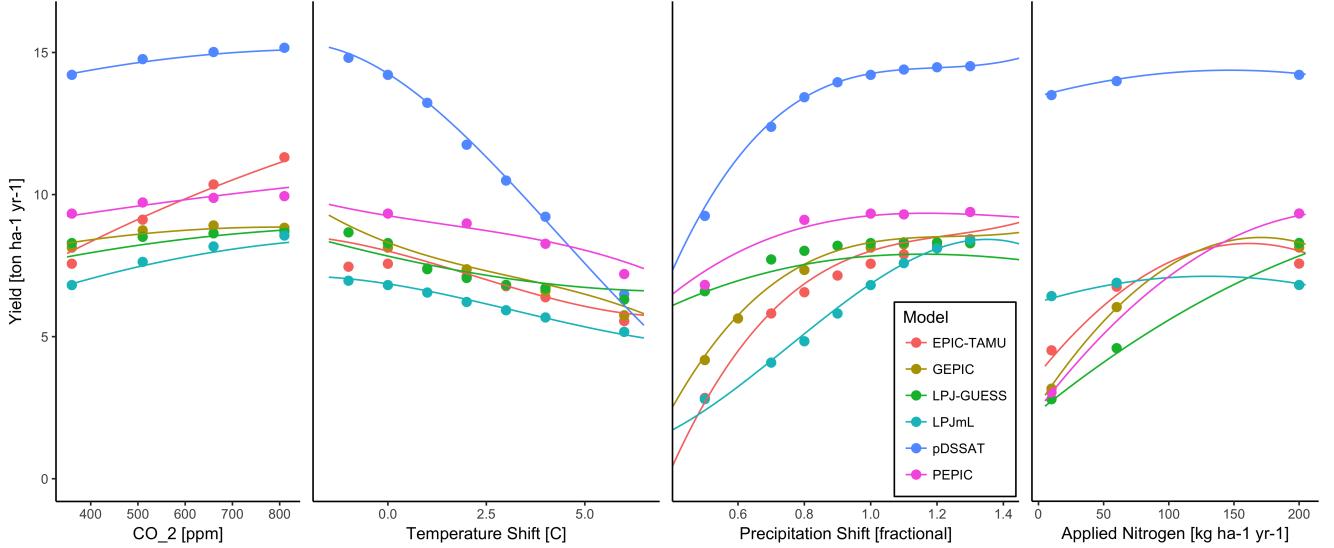


Figure 7: Example illustrating across-model variations in yield response. We show simulations and emulations from six models for rain-fed maize in the same Iowa location shown in Figure 6, with the same plot conventions. Models that did not simulate the nitrogen dimension are omitted for clarity. The inter-model standard deviation is larger for the perturbations in CO₂ and nitrogen than for those in temperature or precipitation illustrating that the sensitivity to weather is better constrained than to management at elevated CO₂. While most model responses can readily be emulated with a simple polynomial, some response surfaces are more complicated and lead to emulation error.

can range from ~5–50%, and nitrogen responses from nearly flat to a 60% drop in the lowest-application simulation. While the nitrogen dimension is important and uncertain, it is also the most problematic to emulate in this work because of its limited sampling. The GGCMI protocol specified only

three nitrogen levels (10, 60 and 200 kg N y⁻¹ ha⁻¹), so a third-order fit would be over-determined but a second-order fit can result in potentially unphysical results. Steep and nonlinear declines in yield with lower nitrogen levels means that some regressions imply a peak in yield between the 100 and 200 kg N

574 $\text{y}^{-1} \text{ ha}^{-1}$ levels. While there may be some reason to believe
 575 over-application of nitrogen at the wrong time in the growing
 576 season could lead to reduced yields, these features are almost
 577 certainly an artifact of under sampling. In addition, the polyno-
 578 mial fit cannot capture the well-documented saturation effect
 579 of nitrogen application (e.g. Ingestad, 1977) as accurately as
 580 would be possible with a non-parametric model.

581 To assess the ability of the polynomial emulation to capture
 582 the behavior of complex process-based models, we evaluate the
 583 normalized emulator error. That is, for each grid cell, model,
 584 and scenario we evaluate the difference between the model yield
 585 and its emulation, normalized by the inter-model standard de-
 586 viation in yield projections. This metric implies that emulation
 587 is generally satisfactory, with several distinct exceptions. Al-
 588 most all model-crop combination emulators have normalized
 589 errors less than one over nearly all currently cultivated hectares
 590 (Figure 8), but some individual model-crop combinations are
 591 problematic (e.g. PROMET for rice and soy, JULES for soy
 592 and winter wheat, Figures S14–S15). Normalized errors for soy
 593 are somewhat higher across all models not because emulator fi-
 594 delity is worse but because models agree more closely on yield
 595 changes for soy than for other crops (see Figure S16, lower-
 596 ing the denominator). Emulator performance often degrades in
 597 geographic locations where crops are not currently cultivated.
 598 Figure 9 shows a CARAIB case as an example, where emulator
 599 performance is satisfactory over cultivated areas for all crops
 600 other than soy, but uncultivated regions show some problematic
 601 areas.

602 It should be noted that this assessment metric is relatively
 603 forgiving. First, each emulation is evaluated against the sim-
 604 ulation actually used to train the emulator. Had we used a spline
 605 interpolation the error would necessarily be zero. Second, the
 606 performance metric scales emulator fidelity not by the magni-
 607 tude of yield changes but by the inter-model spread in those

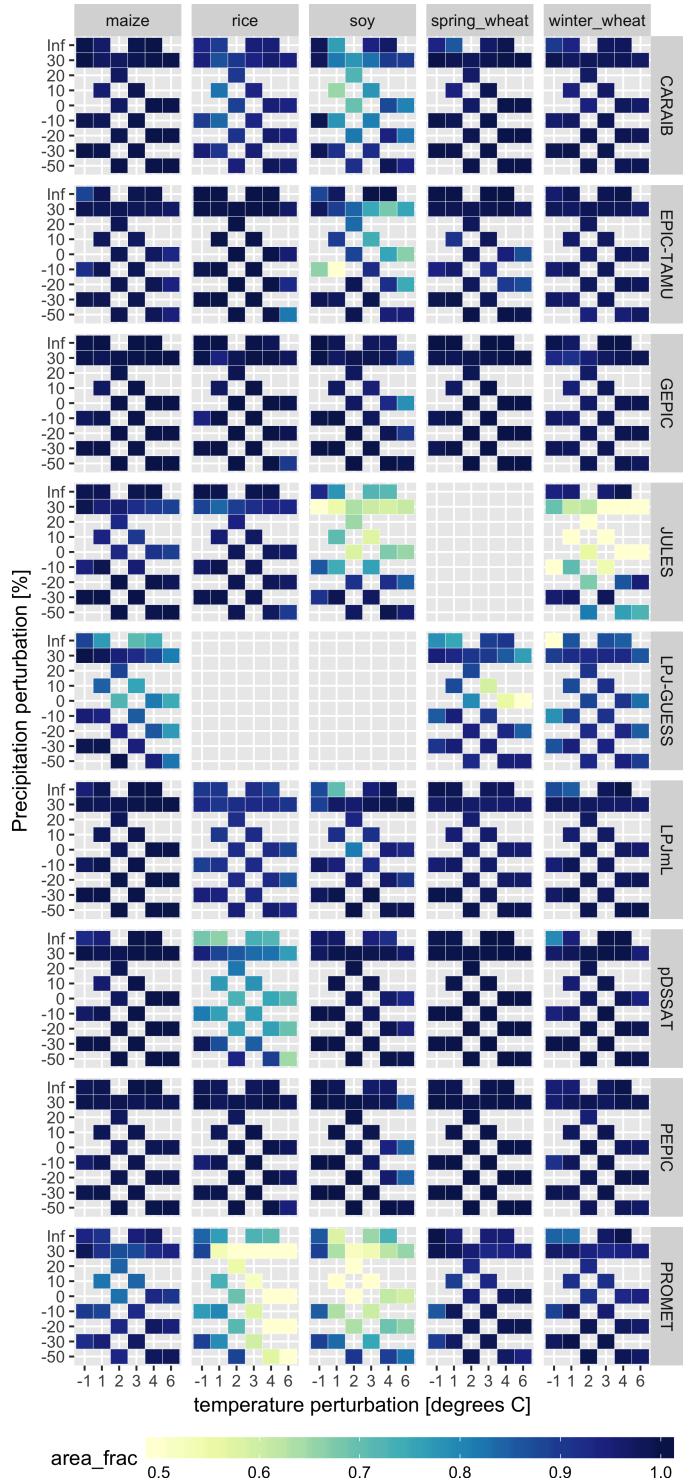


Figure 8: Assessment of emulator performance over currently cultivated areas based on normalized error (Equations 3, 2). We show performance of all 9 models emulated, over all crops and all sampled T and P inputs, but with CO_2 and nitrogen held fixed at baseline values. Large columns are crops and large rows models; squares within are T,P scenario pairs. Colors denote the fraction of currently cultivated hectares with $e < 1$. Of the 756 scenarios with these CO_2 and N values, we consider only those for which all 9 models submitted data. JULES did not simulate spring wheat and LPJ-GUESS did not simulate rice and soy. Emulator performance is generally satisfactory, with some exceptions. Emulator failures (significant areas of poor performance) occur for individual crop-model combinations, with performance generally degrading for hotter and wetter scenarios.

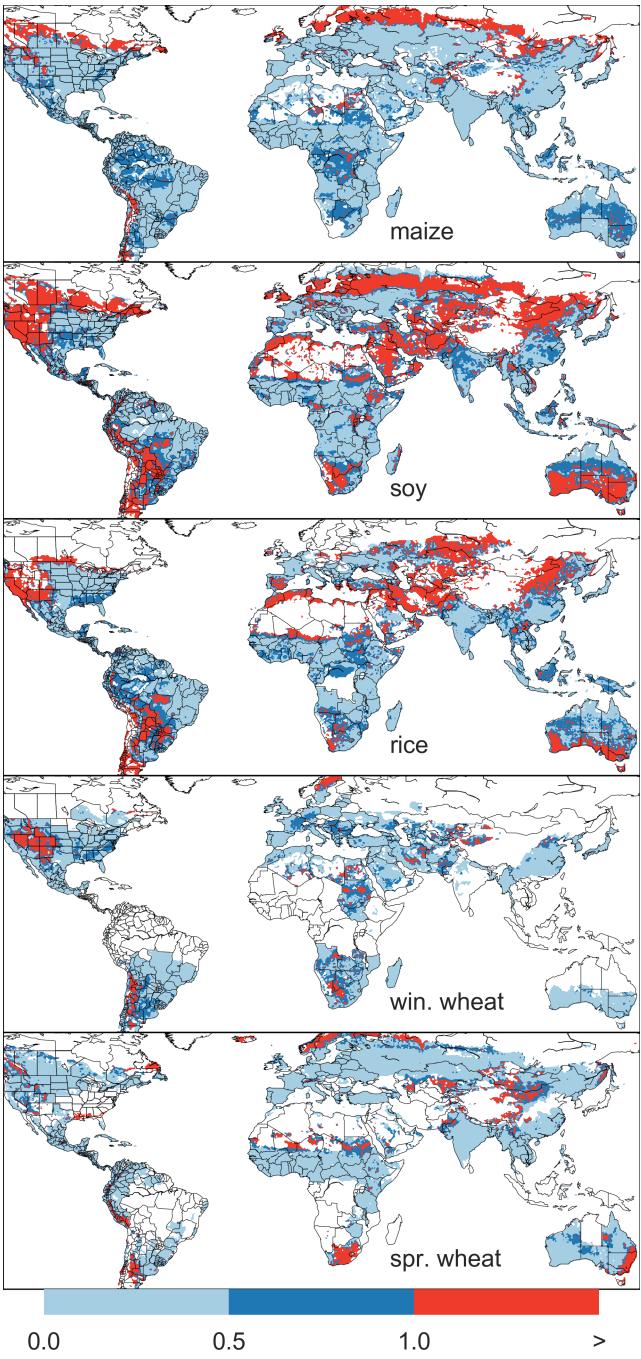


Figure 9: Illustration of our test of emulator performance, applied to the CARAIB model for the T+4 scenario for rain-fed crops. Contour colors indicate the normalized emulator error e , where $e > 1$ means that emulator error exceeds the multi-model standard deviation. White areas are those where crops are not simulated by this model. Models differ in their areas omitted, meaning the number of samples used to calculate the multi-model standard deviation is not spatially consistent in all locations. Emulator performance is generally good relative to model spread in areas where crops are currently cultivated (compare to Figure 1) and in temperate zones in general; emulation issues occur primarily in marginal areas with low yield potentials. For CARAIB, emulation of soy is more problematic, as was also shown in Figure 8.

changes. Where models differ more widely, the standard for emulators becomes less stringent. Because models disagree on the magnitude of CO₂ fertilization, this effect is readily seen when comparing assessments of emulator performance in simulations at baseline CO₂ (Figure 8) with those at higher CO₂ levels (Figure S13). Widening the inter-model spread leads to an apparent increase in emulator skill.

615 3.4. Emulator applications

Because the emulator or “surrogate model” transforms the discrete simulation sample space into a continuous response surface at any geographic scale, it can be used for a variety of applications. Emulators provide an easy way to compare a ensembles of climate or impacts projections. They also provide a means for generating continuous damage functions. As an example, we show a damage function constructed from 4D emulations for aggregated yield at the global scale, for maize on currently cultivated land, with simulated values shown for comparison. (Figure 10; see Figures S16- S19 in the supplemental material for other crops and dimensions.) The emulated values closely match simulations even at this aggregation level. Note that these functions are presented only as examples and do not represent true global projections, because they are developed from simulation data with a uniform temperature shift while increases in global mean temperature should manifest non-uniformly. The global coverage of the GGCMI simulations allows impacts modelers to apply arbitrary geographically-varying climate projections, as well as arbitrary aggregation mask, to develop damage functions for any climate scenario and any geopolitical or geographic level.

4. Conclusions and discussion

The GGCMI Phase II experiment assess sensitivities of process-based crop yield models to changing climate and management inputs, and was designed to allow not only comparison

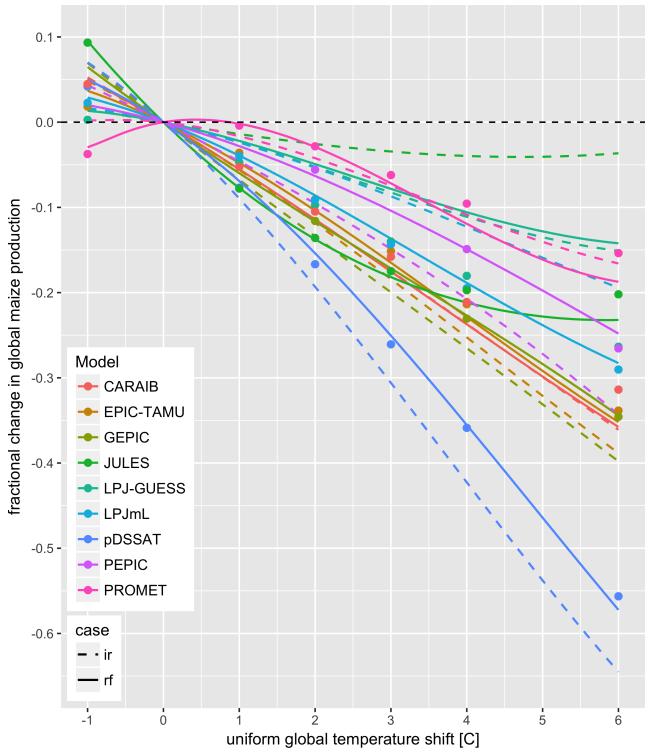


Figure 10: Global emulated damages for maize on currently cultivated lands⁶⁶⁸ for the GGCMI models emulated, for uniform temperature shifts with other inputs held at baseline. (The damage function is created from aggregating up⁶⁶⁹ emulated values at the grid cell level, not from a regression of global mean yields.) Lines are emulations for rain-fed (solid) and irrigated (dashed) crops;⁶⁷⁰ for comparison, dots are the simulated values for the rain-fed case. For most models, irrigated crops show a sharper reduction than do rain-fed because of the⁶⁷¹ locations of cultivated areas: irrigated crops tend to be grown in warmer areas where impacts are more severe for a given temperature shift. (The exceptions⁶⁷² are PROMET, JULES, and LPJmL.) For other crops and scenarios see Figures S16-S19 in the supplemental material.

⁶⁵⁴ mixed, with models performing better for maize and soy than
⁶⁵⁵ for rice and wheat. The value of utilizing multiple models is
⁶⁵⁶ illustrated by the distribution in performance skill across differ-
⁶⁵⁷ ent countries and crops. An end-user of the simulation outputs
⁶⁵⁸ or emulator tool may pick and choose models based on histori-
⁶⁵⁹ cal skill to provide the most faithful temperature and precipita-
⁶⁶⁰ tion response depending on their application. The nitrogen and
⁶⁶¹ CO₂ responses were not validated in this work.

⁶⁶² One counterintuitive result is that irrigated maize shows
⁶⁶³ steeper yield reductions under warming than does rain-fed
⁶⁶⁴ maize when considered only over currently cultivated land. The
⁶⁶⁵ effect is the result of geographic differences in cultivated area.
⁶⁶⁶ In any given location, irrigation increases crop resiliency to
⁶⁶⁷ temperature increase, but irrigated maize is grown in warmer lo-
⁶⁶⁸ cations where the impacts of warming are more severe (Figures
⁶⁶⁹ S5-S6). The same behavior holds for rice and winter wheat,
⁶⁷⁰ but not for soy or spring wheat (Figures S8-S10). Irrigated
⁶⁷¹ wheat and maize are also more sensitive to nitrogen fertiliza-
⁶⁷² tion levels, presumably because growth in rain-fed crops is also
⁶⁷³ water-limited (Figure S19). (Soy as a nitrogen-fixing crop is relatively
⁶⁷⁴ insensitive to nitrogen, and rice is not generally grown in water-
⁶⁷⁵ limited conditions.)

⁶⁴¹ across models but evaluation of complex interactions between⁶⁷⁶
⁶⁴² driving factors (CO₂, temperature, precipitation, and applied
⁶⁷⁷ nitrogen) and identification of geographic shifts in high yield
⁶⁷⁸ potential locations. While the richness of the dataset invites
⁶⁷⁹ further analysis, we show only a selection of insights derived
⁶⁸⁰ from the simulations. Across the major crops, inter-model un-
⁶⁸¹ certainty is greatest for wheat and least for soy. Across factors
⁶⁸² impacting yields, inter-model-uncertainty is largest for CO₂ fer-
⁶⁸³ tilization and nitrogen response effects. Across geographic re-
⁶⁸⁴ gions, inter-model uncertainty is largest in the high latitudes
⁶⁸⁵ where yields may increase, and model projections are most ro-
⁶⁸⁶ bust in low latitudes where yield impacts are largest.
⁶⁸⁷ Model performance when compared to historical data is

We show that emulation of the output of these complex re-
⁶⁸⁸ sponses is possible even with a relatively simple reduced-form
⁶⁸⁹ statistical model and a limited library of simulations. Emula-
⁶⁹⁰ tion therefore offers the opportunity of producing rapid assess-
⁶⁹¹ ments of agricultural impacts for arbitrary climate scenarios in
⁶⁹² a computationally non-intensive way. The resulting tool should
⁶⁹³ aid in impacts assessment, economic studies, and uncertainty
⁶⁹⁴ analyses. Emulator parameter values also provide a useful way
⁶⁹⁵ to compare sensitivities across models to different climate and
⁶⁹⁶ management inputs, and the terms in the polynomial fits offer
⁶⁹⁷ the possibility of physical interpretation of these dependencies
⁶⁹⁸ to some degree.

688 We provide this simulation output dataset for further analysis⁷²² products) or a historical mean emulator (not presented here).
689 by the community as we have only scratched the surface with⁷²³
690 this work. Each simulation run includes year to year variabil-⁷²⁴
691 ity in yields under different climate and management regimes.⁷²⁵
692 Some of the precipitation and temperature space has been lost⁷²⁶
693 due to the aggregation in the time dimension for the emula-⁷²⁷
694 tor presented here (i.e. the + 6 C simulation in the hottest year⁷²⁸
695 of the historical period compared to the coldest historical year,⁷²⁹
696 or precipitation perturbations in the driest historical year etc).⁷³⁰
697 Development of a year-to-year emulator or an emulator at dif-⁷³¹
698 ferent spatial scales may provide useful for some IAM appli-⁷³²
699 cations. More exhaustive analysis of different statistical model
700 specification for emulation will likely provide additional pre-⁷³³
701 dictive skill over the specification provided here. The poten-⁷³⁴
702 tially richest area for further analysis is the interactions be-⁷³⁵
703 tween input variable especially the Nitrogen and CO₂ interac-⁷³⁶
704 tions with weather and with each other. More robust quantifica-⁷³⁷
705 tion of the sensitivity to the input drivers (and there differences⁷³⁸
706 across models), as well as quantification in differences in un-⁷³⁹
707 certainty across input drivers. Adaptation via growing season⁷⁴⁰
708 changes were also simulated and are available in the database,⁷⁴¹
709 though this dimension was not presented or analyzed here.⁷⁴²

710 The emulation approach presented here has some limitations.⁷⁴³
711 Because the GGCMI simulations apply uniform perturbations⁷⁴⁴
712 to historical climate inputs, they do not sample changes in⁷⁴⁵
713 higher order moments. The emulation therefore does not ad-⁷⁴⁶
714 dress the crop yield impacts of potential changes in climate⁷⁴⁷
715 variability. While some information could be extracted from⁷⁴⁸
716 consideration of year-over-year variability, more detailed sim-⁷⁴⁹
717 ulations and analysis are likely necessary to diagnose the im-⁷⁵⁰
718 pact of changes in variance and sub-growing-season tempo-⁷⁵¹
719 ral effects. Additionally, the emulator is intended to provide⁷⁵²
720 the change in yield from a historical mean baseline value and⁷⁵³
721 should be used in conjunction with historical data (or data prod-⁷⁵⁴
722 ucts) or a historical mean emulator (not presented here).
723 The future of food security is one of the larger challenges
724 facing humanity at present. The development (and emulation)
725 of multi-model ensembles such as GGCMI Phase II provides
726 a way to begin to quantify uncertainties in crop responses to
727 a range of potential climate inputs and explore the potential
728 benefits of adaptive responses. Emulation also allow making
729 state-of-the-art simulation results available to a wide research
730 community as simple, computationally tractable tools that can
731 be used by downstream modelers to understand the socioeco-
732 nomic impacts of crop response to climate change.

5. Acknowledgments

We thank Michael Stein and Kevin Schwarzwald, who provided helpful suggestions that contributed to this work. This research was performed as part of the Center for Robust Decision-Making on Climate and Energy Policy (RDCEP) at the University of Chicago, and was supported through a variety of sources. RDCEP is funded by NSF grant #SES-1463644 through the Decision Making Under Uncertainty program. J.F. was supported by the NSF NRT program, grant #DGE-1735359. C.M. was supported by the MACMIT project (01LN1317A) funded through the German Federal Ministry of Education and Research (BMBF). C.F. was supported by the European Research Council Synergy grant #ERC-2013-SynG-610028 Imbalance. P.P.F. and K.W. were supported by the Newton Fund through the Met Office Climate Science for Service Partnership Brazil (CSSP Brazil). A.S. was supported by the Office of Science of the U.S. Department of Energy as part of the Multi-sector Dynamics Research Program Area. Computing resources were provided by the University of Chicago Research Computing Center (RCC). S.O. acknowledges support from the Swedish strong research areas BECC and MERGE together with support from LUCCI (Lund University Centre for studies of Car-

755 bon Cycle and Climate Interactions).

756 6. References

- 757 Angulo, C., Ritter, R., Lock, R., Enders, A., Fronzek, S., & Ewert, F. (2013).
758 Implication of crop model calibration strategies for assessing regional im-
759 pacts of climate change in europe. *Agric. For. Meteorol.*, 170, 32 – 46.
760 Asseng, S., Ewert, F., Martre, P., Ritter, R. P., B. Lobell, D., Cammarano, D.,
761 A. Kimball, B., Ottman, M., W. Wall, G., White, J., Reynolds, M., D. Al-
762 derman, P., Prasad, P. V. V., Aggarwal, P., Anothai, J., Basso, B., Bier-
763 nath, C., Challinor, A., De Sanctis, G., & Zhu, Y. (2015). Rising tempera-
764 tures reduce global wheat production. *Nature Climate Change*, 5, 143–147.
765 doi:10.1038/nclimate2470.
766 Asseng, S., Ewert, F., Rosenzweig, C., Jones, J., Hatfield, J., Ruane, A.,
767 J. Boote, K., Thorburn, P., Ritter, R. P., Cammarano, D., Brisson, N.,
768 Basso, B., Martre, P., Aggarwal, P., Angulo, C., Bertuzzi, P., Biernath,
769 C., Challinor, A., Doltra, J., & Wolf, J. (2013). Uncertainty in simulat-
770 ing wheat yields under climate change. *Nature Climate Change*, 3, 827832.
771 doi:10.1038/nclimate1916.
772 Aulakh, M. S., & Malhi, S. S. (2005). Interactions of Nitrogen with Other
773 Nutrients and Water: Effect on Crop Yield and Quality, Nutrient Use Ef-
774 ficiency, Carbon Sequestration, and Environmental Pollution. *Advances in*
775 *Agronomy*, 86, 341 – 409.
776 Balkovi, J., van der Velde, M., Skalsk, R., Xiong, W., Folberth, C., Khabarov,
777 N., Smirnov, A., Mueller, N. D., & Obersteiner, M. (2014). Global wheat
778 production potentials and management flexibility under the representative
779 concentration pathways. *Global and Planetary Change*, 122, 107 – 121.
780 Blanc, E. (2017). Statistical emulators of maize, rice, soybean and wheat yields
781 from global gridded crop models. *Agricultural and Forest Meteorology*, 236,
782 145 – 161.
783 Blanc, E., & Sultan, B. (2015). Emulating maize yields from global gridded
784 crop models using statistical estimates. *Agricultural and Forest Meteorol-*
785 *ogy*, 214-215, 134 – 147.
786 von Bloh, W., Schaphoff, S., Müller, C., Rolinski, S., Waha, K., & Zaehle, S.
787 (2018). Implementing the Nitrogen cycle into the dynamic global vegeta-
788 tion, hydrology and crop growth model LPJmL (version 5.0). *Geoscientific*
789 *Model Development*, 11, 2789–2812.
790 Castruccio, S., McInerney, D. J., Stein, M. L., Liu Crouch, F., Jacob, R. L.,
791 & Moyer, E. J. (2014). Statistical Emulation of Climate Model Projections
792 Based on Precomputed GCM Runs. *Journal of Climate*, 27, 1829–1844.
793 Challinor, A., Watson, J., Lobell, D., Howden, S., Smith, D., & Chhetri, N.
794 (2014). A meta-analysis of crop yield under climate change and adaptation.
795 *Nature Climate Change*, 4, 287 – 291.
796 Conti, S., Gosling, J. P., Oakley, J. E., & O'Hagan, A. (2009). Gaussian process
797 emulation of dynamic computer codes. *Biometrika*, 96, 663–676.
798 Duncan, W. (1972). SIMCOT: a simulation of cotton growth and yield. In
799 C. Murphy (Ed.), *Proceedings of a Workshop for Modeling Tree Growth,*
800 *Duke University, Durham, North Carolina* (pp. 115–118). Durham, North
801 Carolina.
802 Duncan, W., Loomis, R., Williams, W., & Hanau, R. (1967). A model for
803 simulating photosynthesis in plant communities. *Hilgardia*, (pp. 181–205).
804 Dury, M., Hambuckers, A., Warnant, P., Henrot, A., Favre, E., Ouberdoorn,
805 M., & Fran ois, L. (2011). Responses of European forest ecosystems to
806 21st century climate: assessing changes in interannual variability and fire
807 intensity. *iForest - Biogeosciences and Forestry*, (pp. 82–99).
808 Elliott, J., Kelly, D., Chryssanthacopoulos, J., Glotter, M., Jhunjhnuwala, K.,
809 Best, N., Wilde, M., & Foster, I. (2014). The parallel system for integrating
810 impact models and sectors (pSIMS). *Environmental Modelling and Soft-*
811 *ware*, 62, 509–516.
812 Elliott, J., M ller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K. J.,
813 B chner, M., Foster, I., Glotter, M., Heinke, J., Iizumi, T., Izaurrealde, R. C.,
814 Mueller, N. D., Ray, D. K., Rosenzweig, C., Ruane, A. C., & Sheffield, J.
815 (2015). The Global Gridded Crop Model Intercomparison: data and mod-
816 eling protocols for Phase 1 (v1.0). *Geoscientific Model Development*, 8,
817 261–277.
818 Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J.,
819 & Taylor, K. E. (2016). Overview of the coupled model intercomparison
820 project phase 6 (cmip6) experimental design and organization. *Geoscientific*
821 *Model Development*, 9, 1937–1958.
822 Ferrise, R., Moriondo, M., & Bindi, M. (2011). Probabilistic assessments of cli-
823 mate change impacts on durum wheat in the mediterranean region. *Natural*
824 *Hazards and Earth System Sciences*, 11, 1293–1302.
825 Folberth, C., Gaiser, T., Abbaspour, K. C., Schulin, R., & Yang, H. (2012). Re-
826 gionalization of a large-scale crop growth model for sub-Saharan Africa:
827 Model setup, evaluation, and estimation of maize yields. *Agriculture,*
828 *Ecosystems & Environment*, 151, 21 – 33.
829 Food and Agriculture Organization of the United Nations (2018). FAOSTAT
830 database. URL: <http://www.fao.org/faostat/en/home>.
831 Fronzek, S., Pirttioja, N., Carter, T. R., Bindi, M., Hoffmann, H., Palosuo, T.,
832 Ruiz-Ramos, M., Tao, F., Trnka, M., Acutis, M., Asseng, S., Baranowski, P.,
833 Basso, B., Bodin, P., Buis, S., Cammarano, D., Deligios, P., Destain, M.-F.,
834 Dumont, B., Ewert, F., Ferrise, R., Fran ois, L., Gaiser, T., Hlavinka, P.,
835 Jacquemin, I., Kersebaum, K. C., Kollas, C., Krzyszczak, J., Lorite, I. J.,
836 Minet, J., Minguez, M. I., Montesino, M., Moriondo, M., M ller, C., Nen-
837 del, C.,  zt rk, I., Perego, A., Rodr guez, A., Ruane, A. C., Ruget, F., Sanna,
838 M., Semenov, M. A., Slawinski, C., Strattonovich, P., Supit, I., Waha, K.,

- 839 Wang, E., Wu, L., Zhao, Z., & Rötter, R. P. (2018). Classifying multi-models⁸⁸²
 840 wheat yield impact response surfaces showing sensitivity to temperature and⁸⁸³
 841 precipitation change. *Agricultural Systems*, *159*, 209–224. ⁸⁸⁴
- 842 Glotter, M., Elliott, J., McInerney, D., Best, N., Foster, I., & Moyer, E. J. (2014).⁸⁸⁵
 843 Evaluating the utility of dynamical downscaling in agricultural impacts pro-⁸⁸⁶
 844 jections. *Proceedings of the National Academy of Sciences*, *111*, 8776–8781.⁸⁸⁷
- 845 Glotter, M., Moyer, E., Ruane, A., & Elliott, J. (2015). Evaluating the Sensitiv-⁸⁸⁸
 846 ity of Agricultural Model Performance to Different Climate Inputs. *Journal*⁸⁸⁹
 847 *of Applied Meteorology and Climatology*, *55*, 151113145618001. ⁸⁹⁰
- 848 Hank, T., Bach, H., & Mauser, W. (2015). Using a Remote Sensing-Supported⁸⁹¹
 849 Hydro-Agroecological Model for Field-Scale Simulation of Heterogeneous⁸⁹²
 850 Crop Growth and Yield: Application for Wheat in Central Europe. *Remote*⁸⁹³
 851 *Sensing*, *7*, 3934–3965. ⁸⁹⁴
- 852 He, W., Yang, J., Zhou, W., Drury, C., Yang, X., D. Reynolds, W., Wang, H.,⁸⁹⁵
 853 He, P., & Li, Z.-T. (2016). Sensitivity analysis of crop yields, soil water⁸⁹⁶
 854 contents and nitrogen leaching to precipitation, management practices and⁸⁹⁷
 855 soil hydraulic properties in semi-arid and humid regions of Canada using the⁸⁹⁸
 856 DSSAT model. *Nutrient Cycling in Agroecosystems*, *106*, 201–215. ⁸⁹⁹
- 857 Heady, E. O. (1957). An Econometric Investigation of the Technology of Agri-⁹⁰⁰
 858 cultural Production Functions. *Econometrica*, *25*, 249–268. ⁹⁰¹
- 859 Heady, E. O., & Dillon, J. L. (1961). *Agricultural production functions*. Iowa⁹⁰²
 860 State University Press. ⁹⁰³
- 861 Holden, P., Edwards, N., PH, G., Fraedrich, K., Lunkeit, F., E, K., Labriet,⁹⁰⁴
 862 M., Kanudia, A., & F, B. (2014). Plasim-entsem v1.0: A spatiotemporal⁹⁰⁵
 863 emulator of future climate change for impacts assessment. *Geoscientific*⁹⁰⁶
 864 *Model Development*, *7*, 433–451. doi:10.5194/gmd-7-433-2014. ⁹⁰⁷
- 865 Holzkämper, A., Calanca, P., & Fuhrer, J. (2012). Statistical crop models:⁹⁰⁸
 866 Predicting the effects of temperature and precipitation changes. *Climate*⁹⁰⁹
 867 *Research*, *51*, 11–21. ⁹¹⁰
- 868 Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I.,⁹¹¹
 869 McLean, G., Chenu, K., van Oosterom, E. J., Snow, V., Murphy, C., Moore,⁹¹²
 870 A. D., Brown, H., Whish, J. P., Verrall, S., Fainges, J., Bell, L. W., Peake,⁹¹³
 871 A. S., Poulton, P. L., Hochman, Z., Thorburn, P. J., Gaydon, D. S., Dalgliesh,⁹¹⁴
 872 N. P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp,⁹¹⁵
 873 J., Cichota, R., Vogeler, I., Li, F. Y., Wang, E., Hammer, G. L., Robertson,⁹¹⁶
 874 M. J., Dimes, J. P., Whitbread, A. M., Hunt, J., van Rees, H., McClelland, T.,⁹¹⁷
 875 Carberry, P. S., Hargreaves, J. N., MacLeod, N., McDonald, C., Harsdorff, J.,⁹¹⁸
 876 Wedgwood, S., & Keating, B. A. (2014). APSIM Evolution towards a new⁹¹⁹
 877 generation of agricultural systems simulation. *Environmental Modelling and*⁹²⁰
 878 *Software*, *62*, 327 – 350. ⁹²¹
- 879 Howden, S., & Crimp, S. (2005). Assessing dangerous climate change impacts⁹²²
 880 on australia's wheat industry. *Modelling and Simulation Society of Australia*⁹²³
 881 and New Zealand, (pp. 505–511). ⁹²⁴
- Iizumi, T., Nishimori, M., & Yokozawa, M. (2010). Diagnostics of climate
 model biases in summer temperature and warm-season insolation for the
 simulation of regional paddy rice yield in japan. *Journal of Applied Meteor-
 ology and Climatology*, *49*, 574–591.
- Ingestad, T. (1977). Nitrogen and Plant Growth; Maximum Efficiency of Ni-
 trogen Fertilizers. *Ambio*, *6*, 146–151.
- Izaurralde, R., Williams, J., McGill, W., Rosenberg, N., & Quiroga Jakas, M.
 (2006). Simulating soil C dynamics with EPIC: Model description and test-
 ing against long-term data. *Ecological Modelling*, *192*, 362–384.
- Jagtap, S. S., & Jones, J. W. (2002). Adaptation and evaluation of the
 CROPGRO-soybean model to predict regional yield and production. *Agri-
 culture, Ecosystems & Environment*, *93*, 73 – 85.
- Jones, J., Hoogenboom, G., Porter, C., Boote, K., Batchelor, W., Hunt, L.,
 Wilkens, P., Singh, U., Gijsman, A., & Ritchie, J. (2003). The DSSAT
 cropping system model. *European Journal of Agronomy*, *18*, 235 – 265.
- Jones, J. W., Antle, J. M., Basso, B., Boote, K. J., Conant, R. T., Foster, I.,
 Godfray, H. C. J., Herrero, M., Howitt, R. E., Janssen, S., Keating, B. A.,
 Munoz-Carpena, R., Porter, C. H., Rosenzweig, C., & Wheeler, T. R. (2017).
 Toward a new generation of agricultural system data, models, and knowl-
 edge products: State of agricultural systems science. *Agricultural Systems*,
155, 269 – 288.
- Keating, B., Carberry, P., Hammer, G., Probert, M., Robertson, M., Holzworth,
 D., Huth, N., Hargreaves, J., Meinke, H., Hochman, Z., McLean, G., Ver-
 burg, K., Snow, V., Dimes, J., Silburn, M., Wang, E., Brown, S., Bristow, K.,
 Asseng, S., Chapman, S., McCown, R., Freebairn, D., & Smith, C. (2003).
 An overview of APSIM, a model designed for farming systems simulation.
European Journal of Agronomy, *18*, 267 – 288.
- Lindeskog, M., Arneth, A., Bondeau, A., Waha, K., Seaquist, J., Olin, S., &
 Smith, B. (2013). Implications of accounting for land use in simulations of
 ecosystem carbon cycling in Africa. *Earth System Dynamics*, *4*, 385–407.
- Liu, J., Williams, J. R., Zehnder, A. J., & Yang, H. (2007). GEPIC - modelling
 wheat yield and crop water productivity with high resolution on a global
 scale. *Agricultural Systems*, *94*, 478 – 493.
- Liu, W., Yang, H., Folberth, C., Wang, X., Luo, Q., & Schulin, R. (2016a).
 Global investigation of impacts of PET methods on simulating crop-water
 relations for maize. *Agricultural and Forest Meteorology*, *221*, 164 – 175.
- Liu, W., Yang, H., Liu, J., Azevedo, L. B., Wang, X., Xu, Z., Abbaspour, K. C.,
 & Schulin, R. (2016b). Global assessment of nitrogen losses and trade-offs
 with yields from major crop cultivations. *Science of The Total Environment*,
572, 526 – 537.
- Lobell, D. B., & Burke, M. B. (2010). On the use of statistical models to predict
 crop yield responses to climate change. *Agricultural and Forest Meteorol-
 ogy*, *150*, 1443 – 1452.

- 925 Lobell, D. B., & Field, C. B. (2007). Global scale climate-crop yield relation-968
 926 ships and the impacts of recent warming. *Environmental Research Letters*,⁹⁶⁹
 927 2, 014002. ⁹⁷⁰
- 928 MacKay, D. (1991). Bayesian Interpolation. *Neural Computation*, 4, 415–447.⁹⁷¹
 929 Makowski, D., Asseng, S., Ewert, F., Bassu, S., Durand, J., Martre, P.,⁹⁷²
 930 Adam, M., Aggarwal, P., Angulo, C., Baron, C., Basso, B., Bertuzzi,⁹⁷³
 931 P., Biernath, C., Boogaard, H., Boote, K., Brisson, N., Cammarano,⁹⁷⁴
 932 D., Challinor, A., Conijn, J., & Wolf, J. (2015). Statistical analysis of⁹⁷⁵
 933 large simulated yield datasets for studying climate effects. (p. 1100).⁹⁷⁶
 934 doi:10.13140/RG.2.1.5173.8328. ⁹⁷⁷
- 935 Mauser, W., & Bach, H. (2015). PROMET - Large scale distributed hydrolog-978
 936 ical modelling to study the impact of climate change on the water flows of⁹⁷⁹
 937 mountain watersheds. *Journal of Hydrology*, 376, 362 – 377. ⁹⁸⁰
- 938 Mauser, W., Klepper, G., Zabel, F., Delzeit, R., Hank, T., Putzenlechner, B.,⁹⁸¹
 939 & Calzadilla, A. (2009). Global biomass production potentials exceed ex-⁹⁸²
 940 pected future demand without the need for cropland expansion. *Nature Com-*⁹⁸³
941 munications, 6. ⁹⁸⁴
- 942 McDermid, S., Dileepkumar, G., Murthy, K., Nedumaran, S., Singh, P., Srinivas, C.,⁹⁸⁵
 943 Gangwar, B., Subash, N., Ahmad, A., Zubair, L., & Nissanka, S.⁹⁸⁶
 944 (2015). Integrated assessments of the impacts of climate change on agricul-⁹⁸⁷
 945 ture: An overview of AgMIP regional research in South Asia. *Chapter in:*⁹⁸⁸
946 Handbook of Climate Change and Agroecosystems, (pp. 201–218). ⁹⁸⁹
- 947 Mistry, M. N., Wing, I. S., & De Cian, E. (2017). Simulated vs. empirical⁹⁹⁰
 948 weather responsiveness of crop yields: US evidence and implications for⁹⁹¹
 949 the agricultural impacts of climate change. *Environmental Research Letters*,⁹⁹²
 950 12. ⁹⁹³
- 951 Moore, F. C., Baldos, U., Hertel, T., & Diaz, D. (2017). New science of climate⁹⁹⁴
 952 change impacts on agriculture implies higher social cost of carbon. *Nature*⁹⁹⁵
953 Communications, 8. ⁹⁹⁶
- 954 Mueller, N., Gerber, J., Johnston, M., Ray, D., Ramankutty, N., & Foley, J.⁹⁹⁷
 955 (2012). Closing yield gaps through nutrient and water management. *Nature*,⁹⁹⁸
 956 490, 254–7. ⁹⁹⁹
- 957 Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais,¹⁰⁰⁰
 958 P., Deryng, D., Folberth, C., Glotter, M., Hoek, S., Iizumi, T., Izaurralde,¹⁰⁰¹
 959 R. C., Jones, C., Khabarov, N., Lawrence, P., Liu, W., Olin, S., Pugh, T.,¹⁰⁰²
 960 A. M., Ray, D. K., Reddy, A., Rosenzweig, C., Ruane, A. C., Sakurai, G.,¹⁰⁰³
 961 Schmid, E., Skalsky, R., Song, C. X., Wang, X., de Wit, A., & Yang, H.,¹⁰⁰⁴
 962 (2017). Global gridded crop model evaluation: benchmarking, skills, de-¹⁰⁰⁵
 963 ficiencies and implications. *Geoscientific Model Development*, 10, 1403–¹⁰⁰⁶
 964 1422. ¹⁰⁰⁷
- 965 Nakamura, T., Osaki, M., Koike, T., Hanba, Y. T., Wada, E., & Tadano, T.,¹⁰⁰⁸
 966 (1997). Effect of CO₂ enrichment on carbon and nitrogen interaction in¹⁰⁰⁹
 967 wheat and soybean. *Soil Science and Plant Nutrition*, 43, 789–798. ¹⁰¹⁰
- O'Hagan, A. (2006). Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering & System Safety*, 91, 1290 – 1300.
- Olin, S., Schurges, G., Lindeskog, M., Wårliind, D., Smith, B., Bodin, P., Holmér, J., & Arneth, A. (2015). Modelling the response of yields and tissue C:N to changes in atmospheric CO₂ and N management in the main wheat regions of western europe. *Biogeosciences*, 12, 2489–2515. doi:10.5194/bg-12-2489-2015.
- Osaki, M., Shinano, T., & Tadano, T. (1992). Carbon-nitrogen interaction in field crop production. *Soil Science and Plant Nutrition*, 38, 553–564.
- Osborne, T., Gornall, J., Hooker, J., Williams, K., Wiltshire, A., Betts, R., & Wheeler, T. (2015). JULES-crop: a parametrisation of crops in the Joint UK Land Environment Simulator. *Geoscientific Model Development*, 8, 1139–1155.
- Ostberg, S., Schewe, J., Childers, K., & Frieler, K. (2018). Changes in crop yields and their variability at different levels of global warming. *Earth System Dynamics*, 9, 479–496.
- Oyebamiji, O. K., Edwards, N. R., Holden, P. B., Garthwaite, P. H., Schaphoff, S., & Gerten, D. (2015). Emulating global climate change impacts on crop yields. *Statistical Modelling*, 15, 499–525.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Pas-
 slos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Pirttioja, N., Carter, T., Fronzek, S., Bindi, M., Hoffmann, H., Palosuo, T., Ruiz-Ramos, M., Tao, F., Trnka, M., Acutis, M., Asseng, S., Baranowski, P., Basso, B., Bodin, P., Buis, S., Cammarano, D., Deligios, P., Destain, M., Dumont, B., Ewert, F., Ferrié, R., François, L., Gaiser, T., Hlavinka, P., Jacquemin, I., Kersebaum, K., Kollas, C., Krzyszczak, J., Lorite, I., Minet, J., Minguez, M., Montesino, M., Moriondo, M., Müller, C., Nendel, C., Öztürk, I., Perego, A., Rodríguez, A., Ruane, A., Ruget, F., Sanna, M., Semenov, M., Slawinski, C., Strattonovitch, P., Supit, I., Waha, K., Wang, E., Wu, L., Zhao, Z., & Rötter, R. (2015). Temperature and precipitation effects on wheat yield across a European transect: a crop model ensemble analysis using impact response surfaces. *Climate Research*, 65, 87–105.
- Porter et al. (IPCC) (2014). Food security and food production systems. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. In C. F. et al. (Ed.), *IPCC Fifth Assessment Report* (pp. 485–533). Cambridge, UK: Cambridge University Press.
- Portmann, F., Siebert, S., Bauer, C., & Doell, P. (2008). Global dataset of monthly growing areas of 26 irrigated crops.

- 1011 Portmann, F., Siebert, S., & Doell, 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. *Global Biogeochemical Cycles*, 24, GB1011.
- 1012 1054
1013 1055
1014 1056
1015 1057
1016 1058
1017 1059
1018 1060
1019 1061
1020 1062
1021 1063
1022 1064
1023 1065
1024 1066
1025 1067
1026 1068
1027 1069
1028 1070
1029 1071
1030 1072
1031 1073
1032 1074
1033 1075
1034 1076
1035 1077
1036 1078
1037 1079
1038 1080
1039 1081
1040 1082
1041 1083
1042 1084
1043 1085
1044 1086
1045 1087
1046 1088
1047 1089
1048 1090
1049 1091
1050 1092
1051 1093
1052 1094
1053 1095
Ruane, A. C., Antle, J., Elliott, J., Folberth, C., Hoogenboom, G., Mason-D'Croz, D., Müller, C., Porter, C., Phillips, M. M., Raymundo, R. M., Sands, R., Valdivia, R. O., White, J. W., Wiebe, K., & Rosenzweig, C. (2018). Biophysical and economic implications for agriculture of +1.5° and +2.0°C global warming using AgMIP Coordinated Global and Regional Assessments. *Climate Research*, 76, 17–39.
- Ruane, A. C., Cecil, L. D., Horton, R. M., Gordon, R., McCollum, R., Brown, D., Killough, B., Goldberg, R., Greeley, A. P., & Rosenzweig, C. (2013). Climate change impact uncertainties for maize in panama: Farm information, climate projections, and yield sensitivities. *Agricultural and Forest Meteorology*, 170, 132 – 145.
- Ruane, A. C., Goldberg, R., & Chryssanthacopoulos, J. (2015). Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. *Agric. Forest Meteorol.*, 200, 233–248.
- Ruane, A. C., McDermid, S., Rosenzweig, C., Baigorria, G. A., Jones, J. W., Romero, C. C., & Cecil, L. D. (2014). Carbon-temperature-water change analysis for peanut production under climate change: A prototype for the agmip coordinated climate-crop modeling project (c3mp). *Glob. Change Biol.*, 20, 394–407. doi:10.1111/gcb.12412.
- Rubel, F., & Kottek, M. (2010). Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen-Geiger climate classification. *Meteorologische Zeitschrift*, 19, 135–141.
- Sacks, W. J., Deryng, D., Foley, J. A., & Ramankutty, N. (2010). Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography*, 19, 607–620.
- Schauberger, B., Archontoulis, S., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Elliott, J., Folberth, C., Khabarov, N., Müller, C., A. M. Pugh, T., Rolinski, S., Schaphoff, S., Schmid, E., Wang, X., Schlenker, W., & Frier, K. (2017). Consistent negative response of US crops to high temperatures in observations and crop models. *Nature Communications*, 8, 13931.
- Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106, 15594–15598.
- Snyder, A., Calvin, K. V., Phillips, M., & Ruane, A. C. (2018). A crop yield change emulator for use in gcam and similar models: Persephone v1.0. *Geoscientific Model Development Discussions*, 2018, 1–42.
- Storlie, C. B., Swiler, L. P., Helton, J. C., & Sallaberry, C. J. (2009). Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliability Engineering & System Safety*, 94, 1735 – 1763.
- Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93, 485–498.

- 1097 Tebaldi, C., & Lobell, D. B. (2008). Towards probabilistic projections of cli-
 1098 mate change impacts on global crop yields. *Geophysical Research Letters*,
 1099 35.
- 1100 Valade, A., Ciais, P., Vuichard, N., Viovy, N., Caubel, A., Huth, N., Marin, F., &
 1101 Martin, J. F. (2014). Modeling sugarcane yield with a process-based model
 1102 from site to continental scale: Uncertainties arising from model structure
 1103 and parameter values. *Geoscientific Model Development*, 7, 1225–1245.
- 1104 Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., & Schewe,
 1105 J. (2014). The Inter-Sectoral Impact Model Intercomparison Project
 1106 (ISI-MIP): Project framework. *Proceedings of the National Academy of*
 1107 *Sciences*, 111, 3228–3232.
- 1108 White, J. W., Hoogenboom, G., Kimball, B. A., & Wall, G. W. (2011). Method-
 1109 ologies for simulating impacts of climate change on crop production. *Field*
 1110 *Crops Research*, 124, 357 – 368.
- 1111 Williams, K., Gornall, J., Harper, A., Wiltshire, A., Hemming, D., Quaife, T.,
 1112 Arkebauer, T., & Scoby, D. (2017). Evaluation of JULES-crop performance
 1113 against site observations of irrigated maize from Mead, Nebraska. *Geosci-
 1114 entific Model Development*, 10, 1291–1320.
- 1115 Williams, K. E., & Falloon, P. D. (2015). Sources of interannual yield vari-
 1116 ability in JULES-crop and implications for forcing with seasonal weather
 1117 forecasts. *Geoscientific Model Development*, 8, 3987–3997.
- 1118 de Wit, C. (1957). Transpiration and crop yields. *Verslagen van Land-*
 1119 *bouwkundige Onderzoeken : 64.6.*
- 1120 Wolf, J., & Oijen, M. (2002). Modelling the dependence of european potato
 1121 yields on changes in climate and co2. *Agricultural and Forest Meteorology*,
 1122 112, 217 – 231.
- 1123 Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., Huang, M., Yao,
 1124 Y., Bassu, S., Ciais, P., Durand, J. L., Elliott, J., Ewert, F., Janssens, I. A.,
 1125 Li, T., Lin, E., Liu, Q., Martre, P., Miller, C., Peng, S., Peuelas, J., Ruane,
 1126 A. C., Wallach, D., Wang, T., Wu, D., Liu, Z., Zhu, Y., Zhu, Z., & Asseng,
 1127 S. (2017). Temperature increase reduces global yields of major crops in four
 1128 independent estimates. *Proc. Natl. Acad. Sci.*, 114, 9326–9331.