



**HAL**  
open science

## Multi-wheat-model ensemble responses to interannual climate variability

Alex C. Ruane, Nicholas I. Hudson, Senthold Asseng, Davide Camarrano, Frank Ewert, Pierre Martre, Kenneth J. Boote, Peter J. Thorburn, Pramod K. Aggarwal, Carlos Angulo, et al.

► **To cite this version:**

Alex C. Ruane, Nicholas I. Hudson, Senthold Asseng, Davide Camarrano, Frank Ewert, et al.. Multi-wheat-model ensemble responses to interannual climate variability. *Environmental Modelling and Software*, 2016, 81, pp.86-101. 10.1016/j.envsoft.2016.03.008 . hal-01533879

**HAL Id: hal-01533879**

**<https://hal.science/hal-01533879>**

Submitted on 8 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

## Multi-wheat-model ensemble responses to interannual climate variability

Alex C. Ruane<sup>1</sup>, Nicholas I. Hudson<sup>2</sup>, Senthold Asseng<sup>3</sup>, Davide Camarrano<sup>3,4</sup>, Frank Ewert<sup>5</sup>, Pierre Martre<sup>6,7</sup>, Kenneth J. Boote<sup>3</sup>, Peter J. Thorburn<sup>8</sup>, Pramod K. Aggarwal<sup>9</sup>, Carlos Angulo<sup>5</sup>, Bruno Basso<sup>10</sup>, Patrick Bertuzzi<sup>11</sup>, Christian Biernath<sup>12</sup>, Nadine Brisson<sup>13,14,\*</sup>, Andrew J. Challinor<sup>15,16</sup>, Jordi Doltra<sup>17</sup>, Sebastian Gayler<sup>18</sup>, Richard Goldberg<sup>2</sup>, Robert F. Grant<sup>19</sup>, Lee Heng<sup>20</sup>, Josh Hooker<sup>21</sup>, Leslie A. Hunt<sup>22</sup>, Joachim Ingwersen<sup>18</sup>, Roberto C Izaurralde<sup>23</sup>, Kurt Christian Kersebaum<sup>24</sup>, Soora Naresh Kumar<sup>25</sup>, Christoph Müller<sup>26</sup>, Claas Nendel<sup>24</sup>, Garry O’Leary<sup>27</sup>, Jørgen E. Olesen<sup>28</sup>, Tom M. Osborne<sup>29</sup>, Taru Palosuo<sup>30</sup>, Eckart Priesack<sup>10</sup>, Dominique Ripoche<sup>11</sup>, Reimund P. Rötter<sup>30</sup>, Mikhail A. Semenov<sup>31</sup>, Iurii Shcherbak<sup>32</sup>, Pasquale Steduto<sup>33</sup>, Claudio O. Stöckle<sup>34</sup>, Pierre Stratonovitch<sup>31</sup>, Thilo Streck<sup>18</sup>, Iwan Supit<sup>35</sup>, Fulu Tao<sup>30,36</sup>, Maria Travasso<sup>37</sup>, Katharina Waha<sup>26</sup>, Daniel Wallach<sup>38</sup>, Jeffrey W. White<sup>39</sup> and Joost Wolf<sup>40</sup>

<sup>1</sup>National Aeronautics and Space Administration, Goddard Institute for Space Studies, New York, NY. <sup>2</sup>Columbia University Center for Climate Systems Research, New York, NY. <sup>3</sup>Agricultural & Biological Engineering Department, University of Florida, Gainesville, FL. <sup>4</sup>James Hutton Institute, Invergowrie, Dundee, Scotland, U.K.. <sup>5</sup>Institute of Crop Science and Resource Conservation, Universität Bonn, D-53 115, Germany. <sup>6</sup>National Institute for Agricultural Research (INRA), UMR1095 Genetics, Diversity and Ecophysiology of Cereals (GDEC), F-63 100 Clermont-Ferrand, France. <sup>7</sup>Blaise Pascal University, UMR1095 GDEC, F-63 170 Aubière, France, <sup>8</sup>Commonwealth Scientific and Industrial Research Organization, Agriculture Flagship, Dutton Park QLD 4102, Australia. <sup>9</sup>Consultative Group on International Agricultural Research, Research Program on Climate Change, Agriculture and Food Security, International Water Management Institute, New Delhi 110012, India. <sup>10</sup>Department of Geological Sciences and Kellogg Biological Station, Michigan State University, East Lansing, MI. <sup>11</sup>INRA, US1116 AgroClim, F- 84 914 Avignon, France. <sup>12</sup>Institute of Biochemical Plant Pathology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, D-85 764, Germany. <sup>13</sup>INRA, UMR0211 Agronomie, F-78 750 Thiverval-Grignon, France. <sup>14</sup>AgroParisTech, UMR0211 Agronomie, F-78 750 Thiverval-Grignon, France. <sup>15</sup>Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds LS29JT, UK. <sup>16</sup>CGIAR-ESSP Program on Climate Change, Agriculture and Food Security, International Centre for Tropical Agriculture, A.A. 6713, Cali, Colombia. <sup>17</sup>Cantabrian

1  
2  
3  
4 32 *Agricultural Research and Training Centre, 39600 Muriedas, Spain.* <sup>18</sup>*Institute of Soil Science*  
5  
6 33 *and Land Evaluation, Universität Hohenheim, D-70 599 Stuttgart, Germany.* <sup>19</sup>*Department of*  
7  
8 34 *Renewable Resources, University of Alberta, Edmonton, AB, Canada T6G 2E3.* <sup>20</sup>*International*  
9  
10 35 *Atomic Energy Agency, 1400 Vienna, Austria.* <sup>21</sup>*School of Agriculture, Policy and Development,*  
11  
12 36 *University of Reading, RG6 6AR, United Kingdom.* <sup>22</sup>*Department of Plant Agriculture,*  
13  
14 37 *University of Guelph, Guelph, Ontario, Canada, N1G 2W1.* <sup>23</sup>*Department of Geographical*  
15  
16 38 *Sciences, University of Maryland, College Park, MD 20782.* <sup>24</sup>*Institute of Landscape Systems*  
17  
18 39 *Analysis, Leibniz Centre for Agricultural Landscape Research, D-15 374 Müncheberg,*  
19  
20 40 *Germany.* <sup>25</sup>*Centre for Environment Science and Climate Resilient Agriculture, Indian*  
21  
22 41 *Agricultural Research Institute, New Delhi 110 012, India.* <sup>26</sup>*Potsdam Institute for Climate*  
23  
24 42 *Impact Research, D-14 473 Potsdam, Germany.* <sup>27</sup>*Landscape & Water Sciences, Department of*  
25  
26 43 *Primary Industries, Horsham 3400, Australia.* <sup>28</sup>*Department of Agroecology, Aarhus University,*  
27  
28 44 *8830 Tjele, Denmark.* <sup>29</sup>*National Centre for Atmospheric Science, Department of Meteorology,*  
29  
30 45 *University of Reading, RG6 6BB, United Kingdom.* <sup>30</sup>*Environmental Impacts Group, Natural*  
31  
32 46 *Resources Institute Finland (Luke), FI-01370, Vantaa, Finland.* <sup>31</sup>*Computational and Systems*  
33  
34 47 *Biology Department, Rothamsted Research, Harpenden, Herts, AL5 2JQ, United Kingdom.*  
35  
36 48 <sup>32</sup>*Institute for Future Environments, Queensland University of Technology, Brisbane, QLD 4000,*  
37  
38 49 *Australia,* <sup>33</sup>*Food and Agriculture Organization of the United Nations, Rome, Italy.* <sup>34</sup>*Biological*  
39  
40 50 *Systems Engineering, Washington State University, Pullman, WA 99164-6120.* <sup>35</sup>*Earth System*  
41  
42 51 *Science-Climate Change and Adaptive Land-use and Water Management, Wageningen*  
43  
44 52 *University, 6700AA, The Netherlands.* <sup>36</sup>*Institute of Geographical Sciences and Natural*  
45  
46 53 *Resources Research, Chinese Academy of Science, Beijing 100101, China.* <sup>37</sup>*Institute for Climate*  
47  
48 54 *and Water, INTA-CIRN, 1712 Castelar, Argentina.* <sup>38</sup>*INRA, UMR1248 Agrosystèmes et*  
49  
50 55 *Développement Territorial, F-31 326 Castanet-Tolosan, France.* <sup>39</sup>*Arid-Land Agricultural*  
51  
52 56 *Research Center, USDA-ARS, Maricopa, AZ 85138.* <sup>40</sup>*Plant Production Systems, Wageningen*  
53  
54 57 *University, 6700AA Wageningen, The Netherlands.*  
55  
56  
57  
58  
59  
60  
61

62  
63  
64  
65  
\*Dr Nadine Brisson passed away in 2011 while this work was being carried out.

*Re-Submission Draft – December 18<sup>th</sup>, 2015*

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

62

63 *Keywords: Crop modeling; uncertainty; multi-model ensemble; wheat; AgMIP; climate impacts;*  
64 *temperature; precipitation; interannual variability*

65

66

67

68

69

70

71

72 Corresponding Author's Address

73 Alex C. Ruane

74 Climate Impacts Group, NASA Goddard Institute for Space Studies

75 2880 Broadway

76 New York, NY 10025, USA

77 [Alexander.C.Ruane@nasa.gov](mailto:Alexander.C.Ruane@nasa.gov)

78 *Phone: +1-212-678-5640; Fax: +1-212-678-5645*

79

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

80 **Abstract**

81 We compare 27 wheat models' yield responses to interannual climate variability, analyzed at  
82 locations in Argentina, Australia, India, and The Netherlands as part of the Agricultural Model  
83 Intercomparison and Improvement Project (AgMIP) Wheat Pilot. Each model simulated 1981-  
84 2010 grain yield, and we evaluate results against the interannual variability of growing season  
85 temperature, precipitation, and solar radiation. The amount of information used for calibration  
86 has only a minor effect on most models' climate response, and even small multi-model  
87 ensembles prove beneficial. Wheat model clusters reveal common characteristics of yield  
88 response to climate; however models rarely share the same cluster at all four sites indicating  
89 substantial independence. Only a weak relationship ( $R^2 \leq 0.24$ ) was found between the models'  
90 sensitivities to interannual temperature variability and their response to long-term warming,  
91 suggesting that additional processes differentiate climate change impacts from observed climate  
92 variability analogs and motivating continuing analysis and model development efforts.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

94 **1. Introduction**

95 Process-based crop simulation models have become increasingly prominent in the last several  
96 decades in climate impact research owing to their utility in understanding interactions among  
97 genotype, environment, and management to aid in planning key farm decisions including cultivar  
98 selection, sustainable farm management, and economic planning amidst a variable and changing  
99 climate (e.g., Ewert et al., 2015). In the coming decades climate change is projected to pose  
100 additional and considerable challenges for agriculture and food security around the world (Porter  
101 et al., 2014; Rosenzweig et al., 2014). Process-based crop simulation models have the potential  
102 to provide useful insight into vulnerability, impacts, and adaptation in the agricultural sector by  
103 simulating how cropping systems respond to changing climate, management, and variety choice.  
104 Such gains in insight require high-quality models and better understanding of model  
105 uncertainties for detailed agricultural assessment (Rötter et al., 2011). Although there have been  
106 a large number of studies utilizing crop models to assess climate impacts (Challinor et al.,  
107 2014a), a lack of consistency has made it very difficult to compare results across regions, crops,  
108 models, and climate scenarios (White et al., 2011a). The Agricultural Model Intercomparison  
109 and Improvement Project (AgMIP; Rosenzweig et al., 2013; 2015) was launched in 2010 to  
110 establish a consistent climate-crop-economics modeling framework for agricultural impacts  
111 assessment with an emphasis on multi-model analysis, robust treatment of uncertainty, and  
112 model improvement.

113  
114 A crop model's response to interannual climate variability provides a useful first indicator of  
115 model responses to variation in environmental conditions (Arnold and de Wit, 1976). A  
116 simulation model's ability to capture historical grain yield variability has shown it can serve as a

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

117 sensible basis on which to demonstrate the utility of crop models among stakeholders and  
118 decision-makers (e.g., Dobermann et al., 2000). Considering the effort required in collecting  
119 data and calibrating a crop model for a particular application, previous studies have often relied  
120 upon only a single crop model and limited sets of observational data. This approach overlooks  
121 differences in plausible calibration methodologies as well as biases introduced in the selection of  
122 a single crop model and its parameterization sets; all of which may affect climate sensitivities  
123 (Pirttioja et al., 2015). The final decision-supporting information may therefore be biased  
124 depending on the amount of calibration data available and the crop model selected for  
125 simulations.

126  
127 Here we present an agro-climatic analysis of 27 wheat models that participated in the AgMIP  
128 Wheat Model Intercomparison Pilot (described briefly in the next section and more completely  
129 in the text and supporting materials of Asseng et al., 2013; and Martre et al., 2015), with a focus  
130 on how interannual climate variability affects yield simulations and uncertainties across models.  
131 This is just one of several studies to emerge from the unprecedented Wheat Pilot multi-model  
132 intercomparison and it is intended to contribute to the overall effort by highlighting important  
133 areas for continuing analysis, model improvement, and data collection. As most climate impacts  
134 assessments cannot afford to run all 27 wheat models, for the first time we examine the  
135 consistency of agro-climatic responses across locations, models, and the extent of calibration  
136 information to determine whether a simpler, smaller multi-model assessment may be a suitable  
137 representation of the full AgMIP Wheat Pilot ensemble. The design of the AgMIP Wheat Pilot  
138 also enables a novel comparison of yield responses to interannual climate variability and to mean  
139 climate changes, testing the notion that the response to historical climate variability provides a

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

140 reasonable analog for future climate conditions. The purpose of this analysis is to identify  
141 differences in model behaviors, data limitations, and areas for continuing research and model  
142 improvement.

143

## 144 **2. Materials and Methods**

### 145 *2.1 The AgMIP Wheat Pilot*

146 A total of 27 wheat modeling groups participated in the first phase of the AgMIP Wheat Model  
147 Intercomparison Pilot in order to investigate model performance across a variety of climates,  
148 management regimes, and climate change conditions (focusing on response sensitivity to  
149 temperature and carbon dioxide). This represented the largest multi-model intercomparison of  
150 crop models to date. Major climate change results for grain yields were presented by Asseng et  
151 al. (2013), while Martre et al. (2015) compared model performance across output variables  
152 against field observations. As those studies thoroughly documented the protocols and  
153 participating models of the Wheat Pilot’s first phase, here we summarize the major elements  
154 with an emphasis on factors affecting interannual grain yield variability as simulated at four sites  
155 over the 1981-2010 historical period. Additional work from the Wheat Pilot’s second phase  
156 have focused on response to increases in average temperature (Asseng et al., 2015), and the  
157 models are largely the same as those utilized in phase 1 and analyzed below.

158

#### 159 *2.1.1 Locations*

160 The four locations simulated by participating wheat model groups are shown in **Table 1**, herein  
161 referred to as Argentina (AR), Australia (AU), India (IN), and the Netherlands (NL). Each  
162 location corresponded to a field trial ranked as either “gold” or “platinum” in AgMIP’s field data



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

163 standards (Boote et al., 2015), allowing for detailed model calibration and analysis with high-  
164 quality initial conditions, in-season measurements, phenology, and end-of-season records.  
165 Calibration in this study refers to the process of configuring a crop model for application at a  
166 given site, which typically entails the representation of soil properties, agricultural management,  
167 and coefficients representing the genetic properties of the cultivar planted; the core biophysical  
168 processes are properties of the model developed from extensive experimentation and are  
169 typically not adjusted to match field observations at these sites. These high-quality seasonal data  
170 unfortunately do not correspond to coincident long-term variety trials using the same  
171 management, cultivars, and soils that would be ideal to calibrate interannual variability  
172 (corresponding crop growth observations and long-term variety trials are quite rare, particularly  
173 in developing countries). Even where long-term variety trial data exist (and are publically  
174 available), considerable analysis is needed to attempt a direct comparison with multi-season crop  
175 model simulation given shifts in cultivars every 3-5 years (Piper et al., 1998; Dobermann et al.,  
176 2000; Mavromatis et al., 2001; Singh et al., 2014; Boote et al., 2015). As a result, analysis here  
177 follows many crop modeling studies in utilizing a single-year or short-period (~5 years or less)  
178 field dataset for calibration and then relying on soil properties, plant genetics, and established  
179 model biophysics to determine interannual variability rather than specifically calibrating internal  
180 parameters of response. Palosuo et al. (2011) examined the potential of a smaller multi-model  
181 ensemble to reproduce interannual yield variability of variety trial for wheat having only two  
182 sites with a longer yield series (14+ seasons) but limited data for calibration, finding errors in  
183 each model but much improved statistics for the multi-model ensemble mean. Rötter et al.  
184 (2012) came up with similar results for barley model simulations.

1  
2  
3  
4 186 Daily climate data (maximum and minimum temperatures, solar radiation, precipitation, wind  
5  
6 187 speed, vapor pressure, dew-point temperature, and relative humidity) were compiled from local  
7  
8  
9 188 observations with missing data filled using the NASA Modern Era Retrospective-analysis for  
10  
11 189 Research and Applications (MERRA; Rienecker et al., 2011) and the NASA/GEWEX Solar  
12  
13 190 Radiation Budget (Stackhouse et al., 2011; White et al., 2011b). The Indian site was irrigated  
14  
15  
16 191 according to the field trial applications. The irrigation (date and amount) of the experimental  
17  
18  
19 192 year (Table 1) was used as input to the models for simulating the 30-years historical period  
20  
21 193 although this may not be sufficient for each year. The other sites were rain-fed. Calibration  
22  
23  
24 194 procedures varied from model to model (generally using the field data to detail crop management  
25  
26 195 and soil properties and then configuring cultivar parameters to match growth stage periods). To  
27  
28  
29 196 isolate the climatic signal, the same configuration was used for the historical simulations, future  
30  
31 197 simulations, and the temperature and CO<sub>2</sub> sensitivity tests at each site. The specific calibration  
32  
33  
34 198 approaches were discussed by Challinor et al. (2014b), who found no clear relationship between  
35  
36 199 the number of parameters calibrated and the relative error of harvest index or grain yield. They  
37  
38 200 further noted that this was consistent with compensating errors that can be a benefit of multi-  
39  
40  
41 201 model ensembles but found no evidence of over-tuning in the AgMIP Wheat Pilot.  
42  
43  
44 202

45  
46 203 **Table 1:** Locations simulated in AgMIP Wheat Pilot (for more details see Martre et al., 2015)

Parameter	Location			
	Argentina	Australia	India	Netherlands
Location	Balcarce	Wongan Hills	Delhi	Wageningen
Latitude	37.75°S	30.89°S	28.38°N	51.97°N
Longitude	58.30°W	116.72°E	77.12°E	5.63°E
Cultivar	Oassis	Gamenya	HD2009	Arminda
Irrigated	No	No	Yes (383 mm)	No
N fertilizer (kg N ha <sup>-1</sup> )	120	50	120	160
Planting date	10 August	12 June	23 November	21 October

<b>Anthesis date</b>	23 November	1 October	18 February	20 June
<b>Harvest date</b>	28 December	16 November	3 April	1 August
<b>Year of experiment</b>	1992	1984	1984-1985	1982-1983

204  
205 Additional observations of yields in these regions potentially provide a target for accurate  
206 interannual variability that the models are challenged to match. We therefore examined 1981-  
207 2010 national level yield data from the UN Food and Agricultural Organization  
208 (<http://faostat.fao.org/>), overlapping district-level yields (in Australia; India: Ministry of  
209 Agriculture, Government of India; and the Netherlands: Central Bureau of Statistics, the Hague,  
210 STATLINE), and nearby variety trials (in Argentina: RET, [www.inase.gov.ar](http://www.inase.gov.ar); and the  
211 Netherlands: Central Bureau of Statistics, the Hague, STATLINE) as a point of comparison  
212 against simulated yields. It is not expected that these four modeling locations are precise  
213 representations of the surrounding region; each represents carefully-controlled field trials in one  
214 location within countries characterized by substantial differences in soils, climates, cultivars, and  
215 management practices.

### 216 2.1.2 Wheat Models

217 **Table 2** lists the 27 wheat models that simulated each of the four sites. Details of the processes  
218 and parameter settings that distinguish each of these models are provided in the supplementary  
219 material (particularly Table S2) of Asseng et al. (2013). The AgMIP Wheat Pilot’s first phase  
220 agreed on a policy of model anonymity in the presentation of results, so for the purpose of this  
221 study the models will be referred to only by a number assigned at random. This allowed us to  
222 still determine the range of responses across these models’ native configurations and elucidate  
223 how the selection of a crop model contributes to uncertainty in interannual yield simulations and

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

225 related decisions. The specific mechanisms for each model’s response are being considered in  
226 ongoing analyses and future intercomparison design.

227

### 228 *2.1.3 Types of simulation exercises*

229 Wheat Pilot protocols were designed to investigate whether limitations in data (which hamper  
230 the calibration of crop models in many locations) substantially affect the accuracy of yield  
231 simulation and/or alter the simulated sensitivity to climate variability and climate changes.

232 Participants were therefore instructed to perform simulations in two steps:

233 1) *Low-information simulations*: Weather data, planting, crop emergence, flowering, and  
234 physiological maturity dates, field management information, and soil characteristics and  
235 initial conditions were provided but no information was provided on end-of-season yields  
236 or in-season crop growth and soil water and nitrogen (N) dynamics. This subset of field  
237 experiment data was referred to as “blind test” simulations by Asseng et al. (2013), and  
238 represent the types of data that may be accessible for a large number of locations.

239 2) *High-information simulations*: In addition to the above data modelers were also provided  
240 with in-season growth dynamics from the same years’ field trial, including, leaf area  
241 index (all sites but AU), total above ground biomass and N, root biomass (at IN only),  
242 cumulative evapotranspiration (at AU and IN only), plant available soil water and soil  
243 inorganic N contents within the season (at AU and NL only), and end-of-season grain  
244 yield and protein concentration, and grain density measurements. Plant components  
245 (green leaves, dead leaves, stem, and chaff) biomass and N contents were also available  
246 at NL. This full set of experimental data was referred to as “full calibration” simulations

247 by Asseng et al. (2013) and is equivalent to the more rare gold or platinum standards set  
 248 by Kersebaum et al. (2015) and Boote et al. (2015).

249 Analysis by Asseng et al. (2013) revealed a considerable reduction of biases between field  
 250 observations and yields using the high-information simulations, but noted that both the low- and  
 251 high-information simulations showed a similar response to changes in mean temperature and  
 252 CO<sub>2</sub> concentrations.

253  
 254 **Table 2:** Crop models included in AgMIP Wheat Pilot (in alphabetical order; for more information and details on  
 255 the processes modeled in each model see supplementary materials of Asseng et al., 2013)

Model	Version	Model description and applications	Web address
APES-ACE*	V. 0.9.0.0	(Donatelli et al., 2010; Ewert et al., 2011a)	<a href="http://www.apesimulator.it/default.aspx">http://www.apesimulator.it/default.aspx</a>
APSIM-Nwheats	V.1.55	(Asseng et al., 2004; Asseng et al., 1998; Keating et al., 2003)	<a href="http://www.apsim.info">http://www.apsim.info</a>
APSIM-wheat	V.7.3	(Keating et al., 2003)	<a href="http://www.apsim.info/Wiki/">http://www.apsim.info/Wiki/</a>
AquaCrop*	V.3.1+	(Steduto et al., 2009)	<a href="http://www.fao.org/nr/water/aquacrop.html">http://www.fao.org/nr/water/aquacrop.html</a>
CropSyst	V.3.04.08	(Stockle et al., 2003)	<a href="http://www.bsyse.wsu.edu/CS_Suite/CropSyst/index.html">http://www.bsyse.wsu.edu/CS_Suite/CropSyst/index.html</a>
DSSAT-CERES-Wheat	V.4.0.1.0	(Hoogenboom and White 2003; Jones et al., 2003), (Ritchie et al., 1985)	<a href="http://www.icasa.net/dssat/">http://www.icasa.net/dssat/</a>
DSSAT-CROPSIM-Wheat		(Hunt and Pararajasingham 1995; Jones et al., 2003)	<a href="http://www.icasa.net/dssat/">http://www.icasa.net/dssat/</a>
Ecosys		(Grant et al., 2011)	<a href="https://portal.ales.ualberta.ca/ecosys/">https://portal.ales.ualberta.ca/ecosys/</a>
EPIC wheat		(Kiniry et al., 1995; Williams et al., 1989)	<a href="http://epicapex.brc.tamus.edu/">http://epicapex.brc.tamus.edu/</a>
Expert-N - CERES - wheat	ExpertN 3.0.10 Ceres 2.0	(Biernath et al., 2011; Priesack et al., 2006; Ritchie et al., 1987; Stenger et al., 1999)	<a href="http://www.helmholtz-muenchen.de/en/iboe/expertn/">http://www.helmholtz-muenchen.de/en/iboe/expertn/</a>
Expert-N - GECROS - wheat	ExpertN 3.0.10	(Biernath et al., 2011; Yin and van Laar 2005; Stenger et al., 1999)	<a href="http://www.helmholtz-muenchen.de/en/iboe/expertn/">http://www.helmholtz-muenchen.de/en/iboe/expertn/</a>
Expert-N - SPASS - wheat	ExpertN 3.0.10	(Biernath et al., 2011; Priesack et al., 2006; Stenger et al., 1999; Wang and Engel 2000)	<a href="http://www.helmholtz-muenchen.de/en/iboe/expertn/">http://www.helmholtz-muenchen.de/en/iboe/expertn/</a>
Expert-N - SUCROS – wheat	ExpertN 3.0.10 Sucros2	(Biernath et al., 2011; Goudriaan and Van Laar 1994; Priesack et al., 2006; Stenger et al., 1999)	<a href="http://www.helmholtz-muenchen.de/en/iboe/expertn/">http://www.helmholtz-muenchen.de/en/iboe/expertn/</a>
FASSET	V.2.0	(Berntsen et al., 2003) (Olesen et al., 2002)	<a href="http://www.fasset.dk">http://www.fasset.dk</a>
GLAM-wheat*	V.2	(Challinor et al., 2004; Li et al.,	<a href="http://see-web-">http://see-web-</a>

			2010)	01.leeds.ac.uk/research/icas/climate_change/glam/download_glam.html
HERMES	V.4.26	(Kersebaum 1995; Kersebaum 2007; Kersebaum 2011; Kersebaum and Beblík 2001)		www.zalf.de/en/forschung/institute/lisa/forschung/oekomod/hermes
InfoCrop	V.1	(Aggarwal et al., 2006)		http://www.iari.res.in
LINTUL-4	v.1	(Shibu et al., 2010; Spitters and Schapendonk 1990)		http://models.pps.wur.nl/models
LPJmL*		(Bondeau et al., 2007; Fader et al., 2010; Waha et al., 2012)		http://www.pik-potsdam.de/research/projects/lpjweb
MCWLA-Wheat*	V2.0	(Tao et al., 2009a; Tao and Zhang 2010; Tao et al., 2009b; Tao and Zhang 2011)		---
MONICA	V.1.0	(Nendel et al., 2011)		http://monica.agrosystem-models.com
O'Leary-model	V.7	(Latta and O'Leary 2003; O'Leary and Connor 1996a; b; O'Leary et al., 1985)		Primary documentation for V7 (V3 (O'Leary and Connor 1996a; b), with incremental documentation thereafter.
SALUS	V.1.0	(Basso et al., 2010; Senthilkumar et al., 2009)		www.salusmodel.net
Sirius2010		(Jamieson and Semenov 2000; Jamieson et al., 1998; Lawless et al., 2005; Semenov and Shewry 2011)		http://www.rothamsted.ac.uk/mas-models/sirius.php
SiriusQuality	V.2.0	(Ferrise et al., 2010; He et al., 2011; He et al., 2010; Martre et al., 2006)		http://www1.clermont.inra.fr/siriusquality
STICS	V.1.1	(Brisson et al., 2003; Brisson et al., 1998)		http://www6.paca.inra.fr/stics_eng/
WOFOST*	V.7.1	(van Diepen et al., 1989; Supit and van Diepen, 1994; Boogard et al., 1998)		http://www.wofost.wur.nl

256

257

258 The 1981-2010 historical simulations that form the bulk of these analyses also served as the  
259 historical basis for climate change simulations conducted by each wheat-modeling group. The  
260 same model configurations were therefore forced by the same climate time series and baseline  
261 carbon dioxide concentrations but with historical temperatures adjusted by -3°C, +3°C, +6°C, and  
262 +9°C every day of the year. As initial soil conditions and crop management (including sowing  
263 date and nitrogen fertilizer application) were kept constant over the 30-year period, these  
264 simulations allow for a comparison between model responses to interannual climate variability

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

265 and to mean climate changes. The re-initialization of soil conditions each year reduces the carry-  
266 over effects of multi-year droughts, which reduces overall interannual variability. This is  
267 common in agricultural modeling applications (particularly those that examine future climate  
268 change where the sequence of events is more difficult to project than mean conditions), but  
269 sequential simulations are an important developmental priority for more accurate representation  
270 of extreme events and soil degradation (Basso et al., 2016) and crop rotation effects (Kollas et  
271 al., 2015).

272  
273 *2.2 Performance of Ensemble*

274 Martre et al. (2014) compared grain yield, protein content concentration, and in-season and end-  
275 of-season variables within the 27 wheat model simulations against observations at each of the  
276 four pilot locations. Although some models had the closest match to specific observations,  
277 across all observed variables the 27-model unweighted arithmetic ensemble mean performed  
278 best, in line with earlier findings based on smaller model ensembles even when used to  
279 reproduce interannual yield statistics (Palosuo et al., 2011; Rötter et al., 2012). Thus, while each  
280 wheat model has its own biases and accuracies, the errors across models tended to compensate  
281 and the resulting ensemble had additional value (see also Challinor et al., 2014b). The superior  
282 performance of the ensemble also reflected that wheat models have evolved with enough  
283 independence in approaches to achieve a random distribution of biases for most variables rather  
284 than leading to the emergence of common biases.

285  
286 In light of the superior performance of the 27-member ensemble mean in reproducing field  
287 observations across the four sites (and the lack of long-term historical yield observations at each

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

288 location), for the purposes of this study we utilize the full, 27-model unweighted arithmetic mean  
289 ensemble as the basis for comparison of each model’s climate response.

290

### 291 *2.3 Methods of analysis*

#### 292 *2.3.1. Agro-climatic correlations*

293 As each of the simulations held management constant throughout the 1981-2010 simulation  
294 period and soils were re-initialized each year (with the exception of LPJmL, which did not  
295 reinitialize soil water), interannual yield variability is a result of model responses to climate  
296 factors. Chief among these are precipitation, temperature, and solar radiation, which are likely to  
297 affect crop growth on a number of time scales. Here we focus on the effects of variability in  
298 mean values over the growing season, using Pearson’s correlations against grain yield to  
299 determine key sensitivities within each crop model. Additional variance is likely explained by  
300 climate variables at sub-seasonal time scales (particularly when extreme conditions align with  
301 vulnerable phenological stages), which merits further examination in future studies. Correlation  
302 was chosen as a simple illustration of association between climate and crop model response,  
303 although aspects related to non-linearity and thresholds may not be captured. Future work may  
304 also consider associative metrics such as the probability of detection for extreme events as a way  
305 of isolating important properties of observations and models (Glotter et al., 2016).

306

307 As most studies will not have the luxury of running all 27 wheat models, we investigate the  
308 expected benefit of adding each additional member to a multi-model subset to converge on  
309 behaviors captured by the full 27-model ensemble. Without running the full analysis it is not  
310 possible to know whether the models that are available are among the best or worst for a given



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

311 site’s climate variability response, so we utilize an 80%-exceedance threshold as a practical risk  
312 in simulation design. Results therefore focus on the correlations that would be exceeded by 80%  
313 of the possible combinations for any number of combined models.

314

### 315 *2.3.2 Agro-climatic clustering*

316 We employed the k-means clustering technique to form clusters of wheat models that are  
317 characterized by similar correlations between yield and growing season temperature,  
318 precipitation, and solar radiation (with equal weighting for all). K-means is an iterative process  
319 by which models are regrouped until silhouette values (i.e., similarity between each model and  
320 the other members of its cluster) are maximized. For each location we examined the results with  
321 three, four, and five clusters and visually selected the number that best captured cohesive  
322 groupings in the climate-sensitivity space (this resulted in three clusters in both Argentina and  
323 India and four clusters in both Australia and the Netherlands). Fewer clusters than this grouped  
324 models with substantially different yield sensitivities to climate variability in the same cluster,  
325 while more clusters tended to unnecessarily divide similarly-responsive models. As each model  
326 belongs to a specific cluster at each location, we utilize the frequency that two models appear in  
327 the same clusters across the four sites as a metric of model similarity.

328

## 329 **3. Results and discussion**

### 330 *3.1 Baseline interannual variability*

331 **Figure 1** presents the 1981-2010 yields for the four Wheat Pilot locations from 27 wheat models,  
332 the full model ensemble, and national and regional yields. These high-information simulation  
333 results indicate uncertainty across the model ensemble, although common differences in mean

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

334 yield across the four locations are clear (as discussed by Asseng et al., 2013, and Martre et al.,  
335 2015). Simulations exceed national and regional yields in each location, as wheat models often  
336 do not include the effects of pests, diseases, poor crop management due to labor or equipment  
337 shortages, waterlogging, and other factors that are common on farms outside of experimental  
338 plots. Model results are therefore more representative of yield potential (Evans and Fischer,  
339 1999) than the more complex conditions of a typical farmer's field. The other source of  
340 variation in the gray lines within Figure 1 comes from the less explored interannual variability of  
341 simulated yields, which is the focus of analyses below. Interannual variability is reduced in the  
342 model ensemble, as would be expected from averaging, although noteworthy variations suggest  
343 that there are common behaviors across the crop model responses. Simulated yields (which  
344 examine a single field) are characterized by greater interannual variance compared to the  
345 national and regional level observations, likely because heterogeneities in soils, climate,  
346 cultivars, and management reduces extreme year anomalies when aggregated to scales that may  
347 exceed those of a given extreme event (Ewert et al., 2011b). Only variety trials (in Argentina  
348 and the Netherlands) contain mean and variance of yields that are similar to the simulations,  
349 although differences in management and the varieties cultivated also reduce the utility of these  
350 records as a basis for truth in the comparison of models.

351  
352 Discrepancies between various observational sources and the experimental field simulated by the  
353 wheat models are large enough to caution against an expectation that the models would  
354 reproduce national, regional, or trial-based observational records over the historical period.  
355 These discrepancies are often due to the set up of the simulations from the single field  
356 experiment not representing the diversity of soils, management and cultivars which affected the

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

357 regional and national yield data (but are not documented). Also, yield variability is often driven  
358 by factors other than weather (Ray et al. 2015) and models that are driven by variations in  
359 weather only are bound to not reproduce observational records. As noted above, we therefore  
360 turn to the High-information ensemble average (dark line in Figure 1) as the standard for the  
361 individual crop models given its superior performance in producing the full range of field  
362 observations (Martre et al., 2015). The ensemble also reduces interannual variability through the  
363 averaging of multiple models' potentially uncorrelated anomalies.

364

### 365 *3.2 Effect of calibration on climate sensitivity*

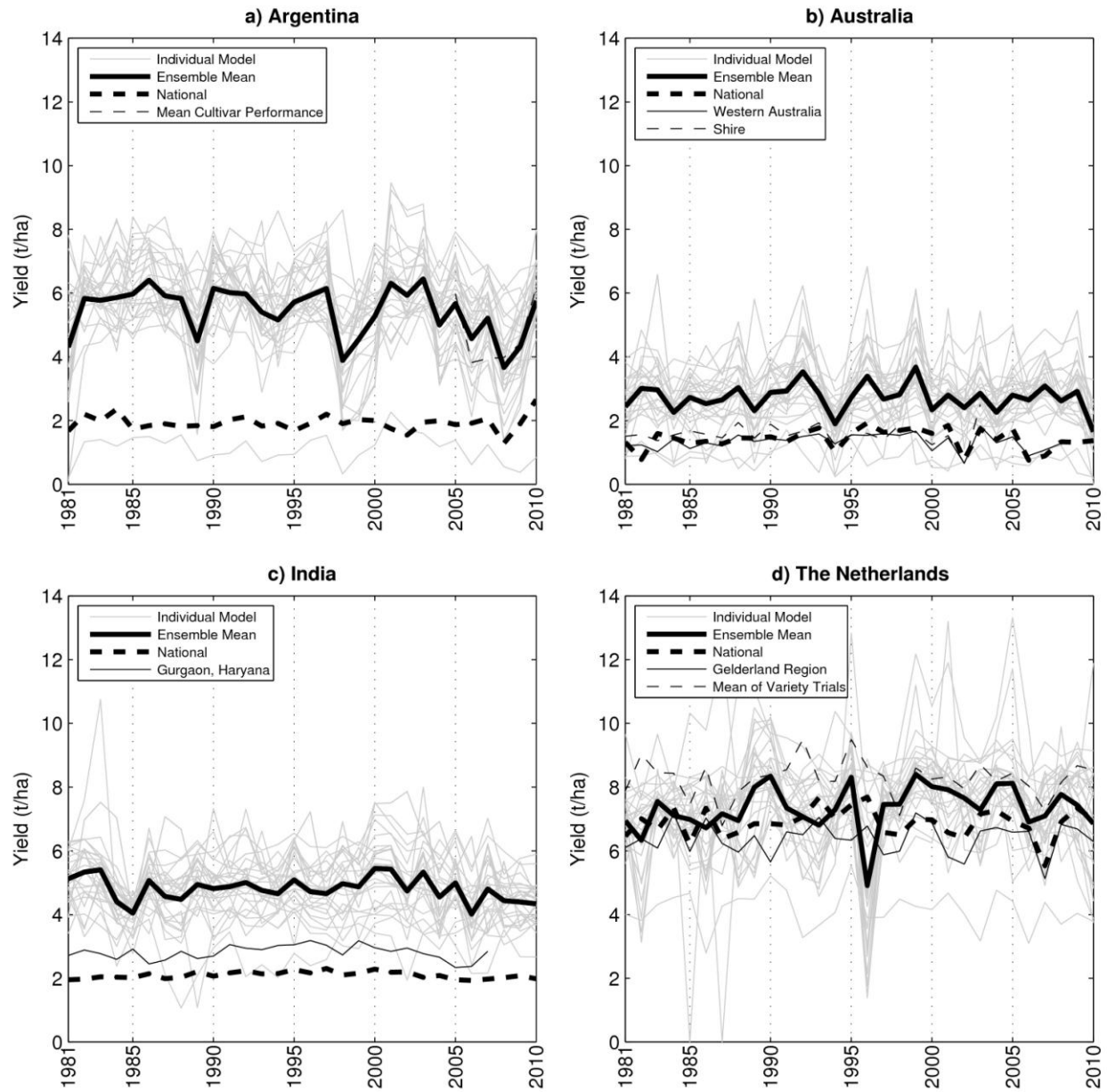
366 The Wheat Pilot's protocol for Low-information and High-information experiments provides a  
367 useful examination of the ways in which model calibration has the potential to affect the  
368 resulting response to climate variability. **Figure 2** illustrates this sensitivity to calibration  
369 information via the correlation of each individual model's low-information results with the full  
370 ensemble of Low-information simulations (LL), the correlation of each model's Low-  
371 information result with the full ensemble of High-information simulations (LH), and the  
372 correlation of each model's High-information results with the full ensemble of High-information  
373 simulations (HH).

374

375 Correlations do not change dramatically between the Low- and High-information simulations for  
376 the vast majority of wheat models at each of the four locations. The exceptions feature both  
377 substantial improvements (e.g., Model #25 in Argentina) and declines (e.g., Model #10 in  
378 Australia) in correlations as additional information is provided. In these cases calibration to  
379 cultivars, soil conditions, or other internal parameters may have improved the experimental

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

380 year's results but also affected climate sensitivity via shifts in the resilience to heat, water, and/or  
381 frost stresses. Effects of calibration strategy on simulations of climate change impact were also  
382 examined by Challinor et al. (2014b) and for simulations of crops across Europe (Angulo et al.,  
383 2013). The relative lack of different sensitivities between the Low- and High-information  
384 simulations could also be explained by the fact that each was simulated by the same model  
385 experts for a given model, and that additional data provided for the High-information  
386  
387



388  
 389 **Figure 1:** Historical period grain yields for a) Argentina, b) Australia, c) India, and d) The Netherlands, including  
 390 the individual crop models at single simulation locations (gray lines), multi-model ensemble mean (black solid line),  
 391 and observations from national, regional, and local field trial data. Linear trends were removed from observational  
 392 data at all but the Argentinian site (which had no significant trend). Modeled yields are the result of the high-  
 393 information calibration simulations.

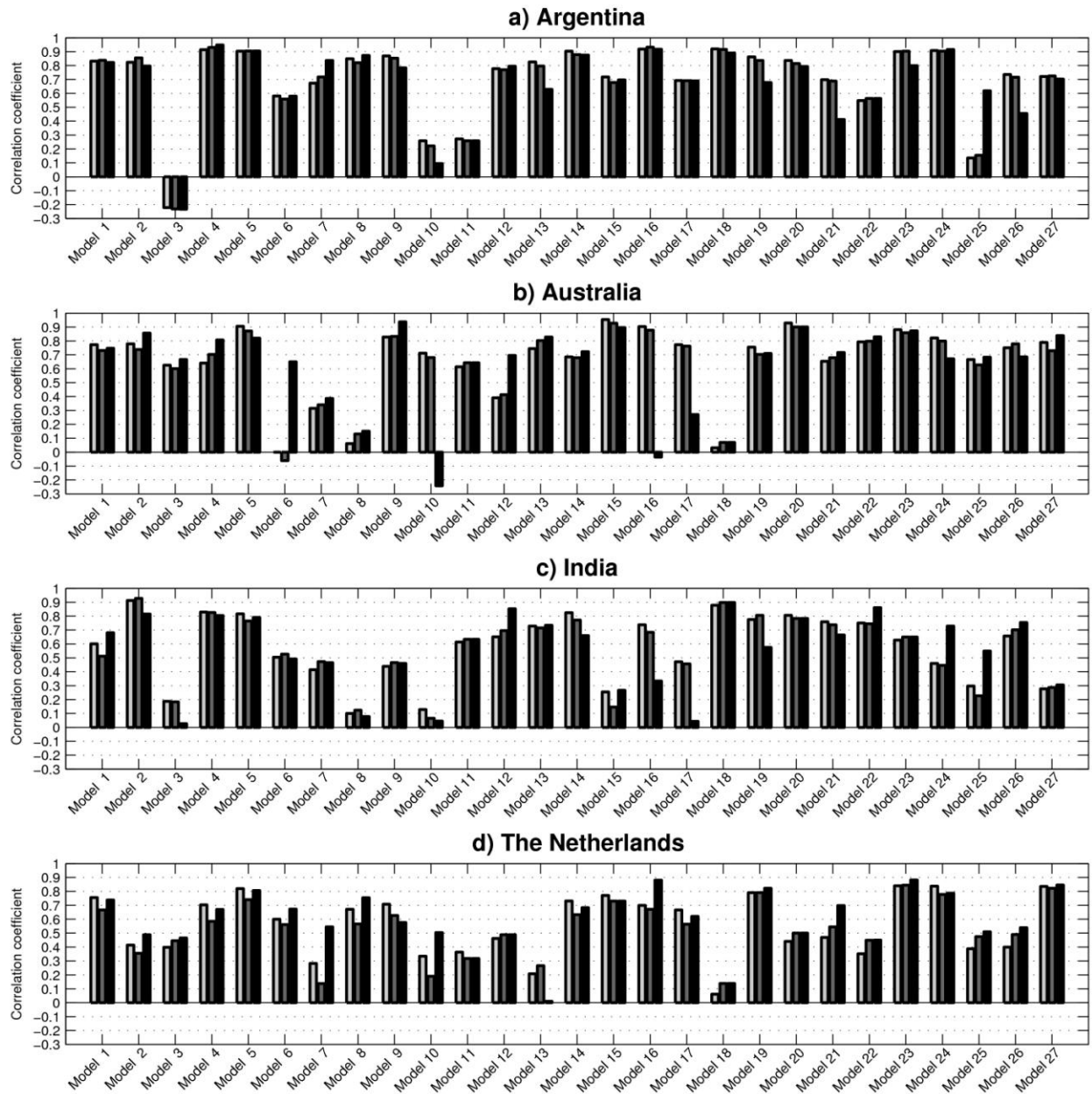
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

395 simulations were mostly limited to details on the crop itself. Additional information about the  
396 soil environment, in particular, would have potentially altered the sensitivity to interannual  
397 rainfall anomalies.

398

399 A comparison between the LL and HH correlations indicates that most models have the same  
400 relationship with the full ensemble regardless of the level of calibration information. Where LH  
401 and HH correlations are similar for a given model there is little benefit from additional  
402 calibration in terms of interannual climate response, as the Low-information results perform just  
403 as well as the High-information results against the High-information ensemble standard. HH  
404 correlations are at least higher than LL correlations in the majority of cases, suggesting that  
405 additional calibration information does tighten the spread of models around the ensemble mean  
406 and thus improve the performance of several models. This benefit is blurred by the likelihood  
407 that the fully-calibrated set of models would be expected to have closer agreement among  
408 members; however, it is important to note that calibration data at each site were only provided  
409 for a single year, making it impossible to directly calibrate the interannual variability examined  
410 here. This is a typical limitation for crop model simulations, as there are few long-term field  
411 trials that would allow full calibration of interannual variability. Also calibration in many cases  
412 focuses on minimizing error between modelled and observed results for the calibration dataset,  
413 which may have little influence on model responses to variation in environmental conditions that  
414 may be controlled by model structure and parameters other than those in focus for the  
415 calibration. The remainder of this study will focus on the High-information simulation sets, as  
416 these are likely to be of highest fidelity. Agro-climatological mechanisms at the root of these  
417 correlations are explored in Section 3.4 below.

418



419  
 420 **Figure 2:** Single model run correlations against ensemble mean during 1981-2010 for (a) Argentina; (b) Australia;  
 421 (c) India; and (d) The Netherlands. The correlation between the Low-information model runs and the Low-  
 422 information ensemble mean (LL) is displayed in light gray, the correlation between the Low-Information model runs  
 423 and the High-information ensemble mean (LH) is displayed in dark gray, and the correlation between the High-  
 424 information model runs and the High-information ensemble mean (HH) is displayed in black.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

425

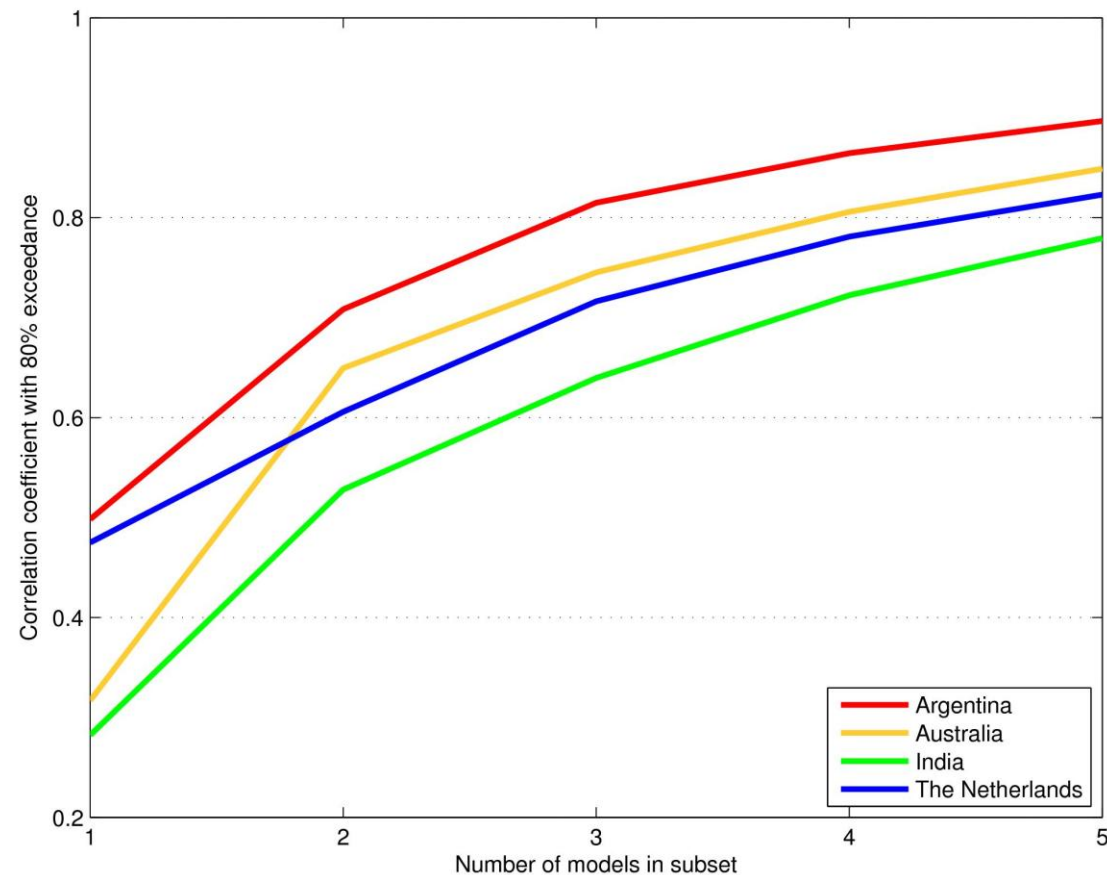
### 426 *3.3 Benefit of multi-model ensemble*

427 The 27-model community approach of the AgMIP Wheat Pilot is not possible in the vast  
428 majority of crop model applications. Instead, what is needed is prior information that aids in the  
429 construction of a practical subset of models with a high likelihood of representing the larger  
430 ensemble. Beginning on the left-hand side of **Figure 3** (representing the use of a randomly  
431 selected single model), the plotted value represents the Pearson's correlation (against the full  
432 High-information ensemble) that would be exceeded by 80% of the individual models. This  
433 value is highest for Argentina (where 80% of the models exceed  $r = 0.50$ ) and lowest for India ( $r$   
434  $= 0.28$ ). Introducing a second model results in  $(27*26)/2=351$  possible combinations, but 80%  
435 of them have a correlation of at least  $r = 0.71$  in Argentina and  $r = 0.53$  in India. Across the four  
436 sites, the benefit of adding a second model to a climate variability analysis is therefore an  
437 increase of +0.23 in its likely correlation with the full ensemble, with gains highest in Australia  
438 (+0.33) and lowest in the Netherlands (+0.13). Adding a third model also substantially increases  
439 the 80%-likely correlation, although the average increase is reduced (+0.11). The additions of a  
440 fourth and fifth model (increasing correlations by an average of 0.06 and 0.04, respectively) to  
441 the subset are also beneficial and lead to very high correlations, but the increases begin to be  
442 small in comparison to the effort likely required to calibrate an additional model (and collaborate  
443 with an additional modeling group) for the effort.

444

445 Efforts to include a second and third model therefore provide substantial benefit to climate  
446 variability simulations; however, investment in including additional models has a diminishing  
447 return. These results suggest a benefit at smaller subsets to account for interannual climate





**Figure 3:** Improvement in correlations with each additional model within a multi-model subset of the full ensemble. For each number of models included in the subset  $N$ , the value shown represents Pearson’s correlation coefficient between the subset’s mean yield and the full ensemble’s mean yield and that would be exceeded 80% of the time given a random selection of  $N$  models from the full set of 27 wheat models. Simulations were performed at single locations in each country (see Table 1) after calibration with High information, and all possible combinations of  $N$  models were tested.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

461

### 3.4. Agro-climatic Sensitivity

463

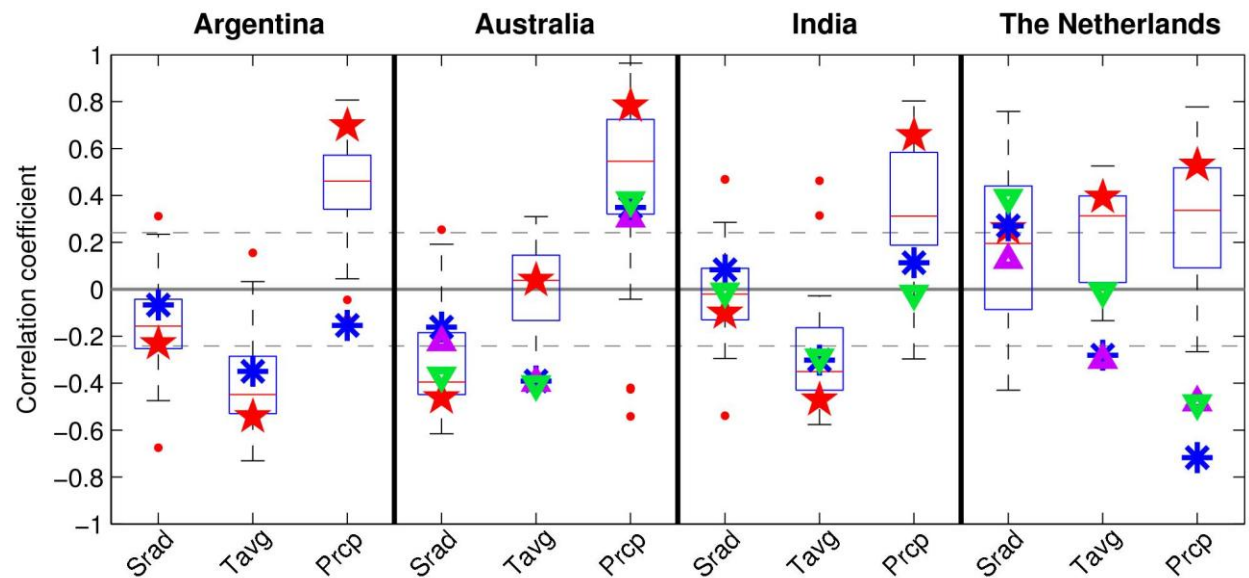
Correlations of the 1981-2010 modeled grain yields and observed grain yields with mean growing season solar radiation, temperature, and precipitation are shown in **Figure 4** across the four locations. In Argentina simulated grain yields are positively correlated with wet seasons in all but one model, with more than 75% of the models demonstrating significant correlations. A strong sensitivity to rainfall anomalies is also seen in the cultivar trials; however, national grain yields are not significantly correlated with the precipitation at Balcarce, Argentina, as the wheat area covers a much larger region. The simulations and cultivar trials agree that lower temperatures significantly favor grain yields, with even the national grain yields following suit as warm and cooler seasons tend to spread more widely than the precipitation anomalies. At all sites, for both temperature and precipitation, the magnitude of the ensemble average's correlation is substantially higher than that of the median model; indicating that precipitation and temperature sensitivities are a unifying factor describing grain yield across the model members. Solar radiation variability is not significantly correlated for the bulk of models.

476

477

The Australian location is characterized by an even stronger sensitivity to rainfall. This site is also significantly sensitive to solar radiation anomalies, with negative correlations suggesting interdependence as cloudier seasons correspond with wetter conditions. National and regional yields are less responsive to precipitation anomalies and are governed more by temperature, as temperature anomalies may be widespread while droughts in the east are often offset by wetter conditions in the west.

483



**Figure 4:** Box-and-whiskers plots of Pearson's correlation coefficients between the 27 wheat models' 1981-2010 simulated grain yields at single locations in each country and corresponding growing season mean solar radiation (Srad), average temperature (Tavg) and precipitation (Prctp). The median of the model simulations is marked by the red line, the box contains the middle two quartiles (from 25% to 75%), and the whiskers extend to the most extreme data points of the simulations that are not considered outliers (displayed as red dots). The correlation of the ensemble performance (red star), national observations (blue asterisk), regional observations (magenta triangles; where available), and the mean of other field trial results or local observations (green triangles) over the years data were available are also presented (as in Figure 1). Dashed lines indicate thresholds for correlations that are significant at the 90<sup>th</sup> percentile (*t*-test).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

501 that are favorable for irrigated wheat growth. Cool seasons here are favorable for wheat  
502 production, and solar radiation correlations are not significant. National level correlations with  
503 the Delhi weather series are understandably weaker for all variables, as heterogeneous climate  
504 across India's wheat-growing regions reduces the prominence of anomalies and results in  
505 insignificant correlations in all but average temperature.

506  
507 Wheat at the Netherlands site follows a different agro-climatic pattern from that at the other three  
508 sites. Warm seasons are positively correlated with yields in the bulk of models, suggesting a  
509 growing degree day limitation. Simulations and observations also suggest a radiation limitation  
510 at this high latitude, with sunnier seasons (and the associated temperature and rainfall patterns)  
511 favoring higher yields. The field site is notably different from the regional and national level  
512 observations in that the aggregated observations are either not correlated with temperature or  
513 suggest that yields favor cooler temperatures. The models also indicate stronger yields in wet  
514 years, while observations indicate better production during drier seasons. This likely comes  
515 from the fact that local and regional management of shallow groundwater tables in this region  
516 helps control against water stress but this management is not considered in the models at the test  
517 site. Contrary to the models' perception of drought, elevated regional yields are recorded in dry  
518 seasons as higher solar radiation and groundwater provisions increase yield potential (Asseng et  
519 al., 2000).

520  
521 *3.5. Clusters of agro-climatic response*

522 **Figure 5** shows each of the 27 wheat models as plotted on a three-dimensional space of  
523 temperature, precipitation, and solar radiation correlations with that model's grain yield. Models

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

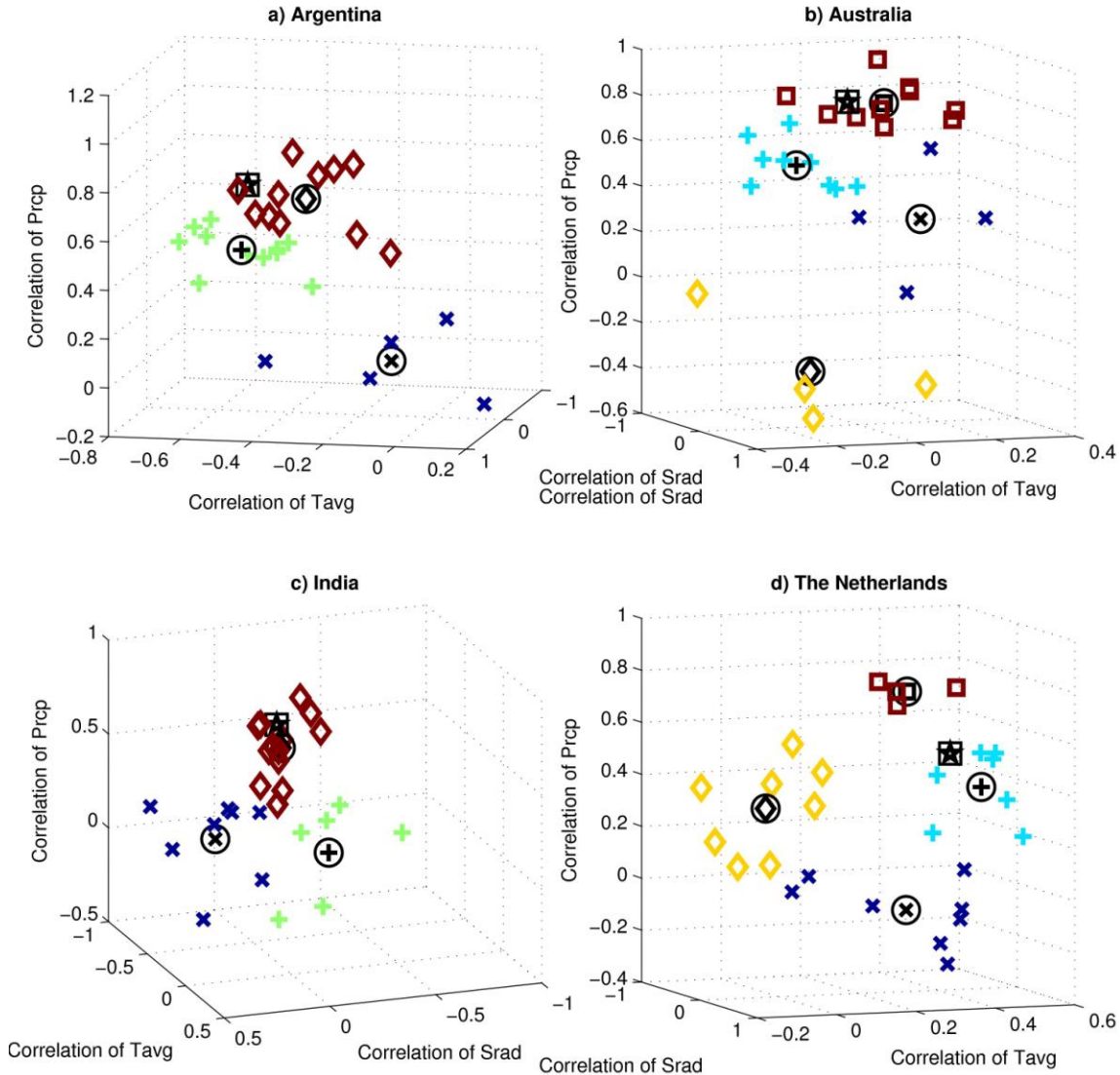
524 falling in the same agro-climatic cluster are represented with a common symbol and color. The  
525 full ensemble average and cluster averages do not fall as an average of the individual model  
526 members' correlations as the ensemble averaging reduces individual models' yearly anomalies to  
527 produce a unique time series. The results illustrate that the model spread is not randomly  
528 distributed in the agro-climatic sensitivity space, but rather distinct families of responses are  
529 evident. Several clusters also correspond much more closely with the full ensemble average  
530 responses.

531

532 **Figure 6** shows the spread of model correlations within each cluster as well as the cluster  
533 ensemble correlations against the full 27-model ensemble's interannual yield variability. One or  
534 two clusters at each location demonstrate substantially better coherence to the ensemble average  
535 than the others. Even within a given cluster there are substantial differences in correlation  
536 between individual models and the ensemble average; particularly among clusters that are  
537 furthest from the ensemble average sensitivities (e.g., the "x" cluster in Argentina or the diamond  
538 cluster in Australia). The ensemble average for each cluster is also a marked improvement on  
539 the median model within that cluster, although occasionally there is one model that outperforms  
540 even the cluster mean.

541

542



**Figure 5:** Clusters of the 27 wheat model simulations (cluster membership denoted by shape of symbols), the ensemble average, and observational data according to their grain yield correlation coefficients versus mean growing season solar radiation (Srad), average temperature (Tavg), and precipitation (Prpc) from 1981-2010 for single locations in (a) Argentina; (b) Australia; (c) India; and (d) The Netherlands. The correlation coefficients of the ensemble yield performance (boxed star) and the centroids of the clusters (corresponding symbols with circles) are also presented. Note that the perspective is rotated and axes limits adjusted in each panel in order to best visualize the differences in the model clusters.

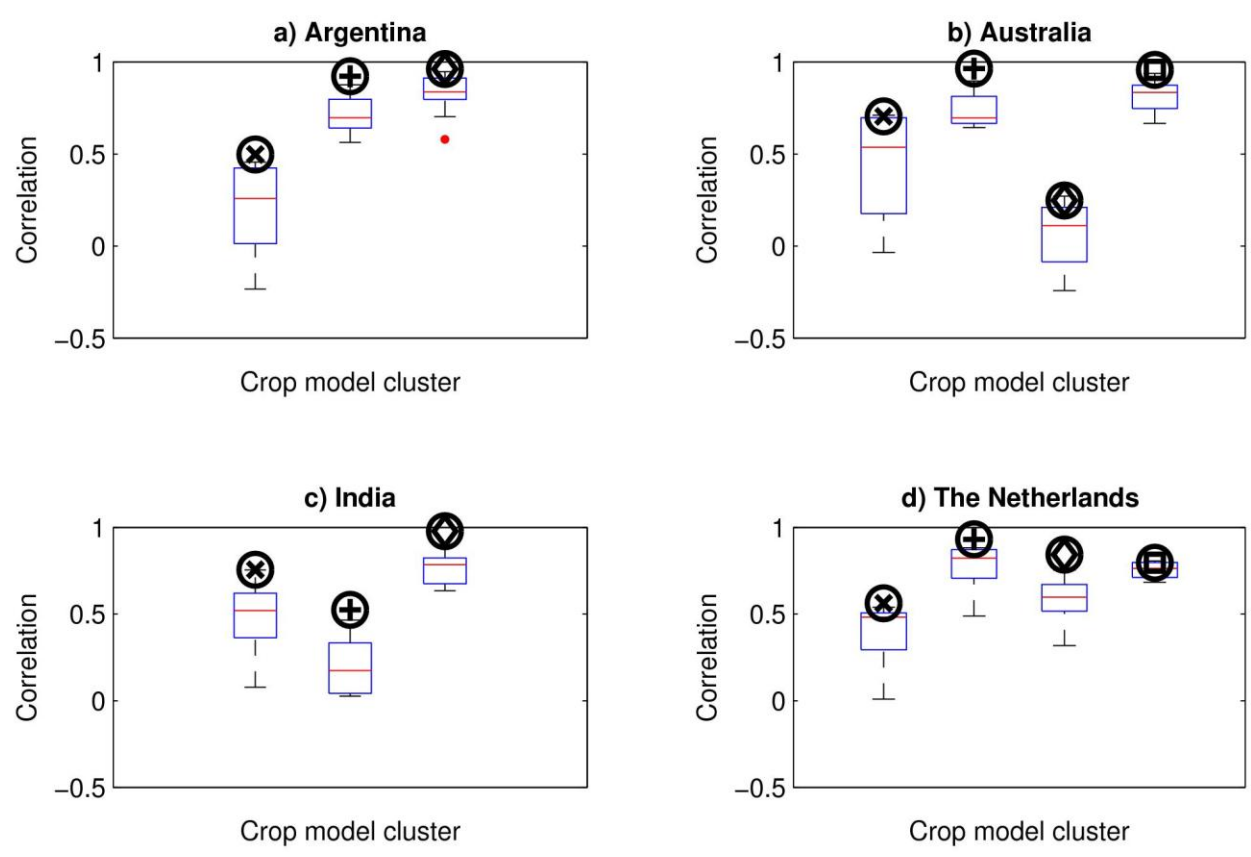
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

553 Despite the fact that many of these wheat models have common heritage in pioneering crop  
554 modeling groups and approaches developed in the last 30 years, only two pairs of models (#1/#5  
555 and #20/#22 from Figure 2; making <0.3% of possible combinations and thus potentially just a  
556 coincidence) fall in the same agro-climatic cluster at all four Pilot locations. 7% of model pairs  
557 fall in the same cluster at three of the four sites, while 24% of model pairs are never in the same  
558 cluster. The remaining 69% of model pairs share one or two clusters, which would be expected  
559 for independent models. No individual model stands out as being particularly divergent from the  
560 others, as each model has at least three other models that never appear in the same cluster, and at  
561 least four models that fall in the same cluster for two or more sites. Only one model falls into the  
562 highest-correlating cluster at all four locations, and likewise only a single model always falls into  
563 the lowest-correlating cluster. In total 15 different models are included in the lowest-correlating  
564 cluster for at least one site, and 21 different models are part of the highest-correlating cluster at  
565 least once. This independence likely contributes to the strength of the full ensemble, as more  
566 independent models are less likely to share common response biases. Model similarities and  
567 differences from site to site also cautions against assuming that performance of a given model at  
568 a limited number of sites is indicative of its likely performance at a new site. The high  
569 sensitivity of the models' response to climate variability demonstrates high sensitivity to  
570 location, representing different growing environments. Results suggest that there is little basis  
571 on which to categorize groups of models based upon expected commonalities in climate  
572 variability response, as these responses show high sensitivity to location rather than models  
573 imposing the same response to all sites.

574  
575 We created subsets of models with the rule that only one model could be drawn from each

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

576 cluster to test the hypothesis that diverse model combinations would more efficiently capture  
577 responses of the full ensemble than would a random combination of wheat models. However,  
578 performance of these subsets was not significantly different from the random subsets tested in  
579 Section 3.3 above. Selecting more diverse models via cluster analysis is therefore not an  
580



581  
582 **Figure 6:** Correlations between simulated grain yield by the wheat models against the 27-member ensemble average  
583 series of interannual grain yields for single locations in (a) Argentina; (b) Australia; (c) India; and (d) the  
584 Netherlands. The correlations of the cluster ensembles are shown in the dark black symbol above the box-and-  
585 whiskers distribution of individual models within that cluster (corresponding to the symbols from Figure 5).

586  
587 effective strategy for creating multi-model subsets for new studies, although the construction of  
588 subsets based upon model structure and parameter sets (rather than response characteristics)



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

589 merits further study. Additional work may also explore agro-climatic responses in perturbed  
590 physics ensembles as an alternative to multi-model ensembles (PPEs and MMEs, respectively;  
591 Wallach et al., 2015).

592  
593

### 594 *3.6. Relationship between interannual and climatological temperature sensitivities*

595 While the above analyses focused on the ways in which simulated grain yields are sensitive to  
596 interannual variability in temperature, rainfall, and solar radiation, the temperature sensitivity  
597 tests (-3°C, +3°C, +6°C, and +9°C) isolate the effect of mean changes in temperature. Popular  
598 impressions of climate change impacts are often based upon *temporal proxies*, or the assumption  
599 that an *x*-degree warmer mean climate at a given location would have grain yields similar to the  
600 yields observed in that location in past years when an *x*-degree anomaly occurred. Empirical  
601 models based upon historical regressions are often premised on such an assumption, although  
602 developed to a greater extent (e.g., Lobell and Burke, 2010). This is indeed a logical hypothesis  
603 as one would expect that a crop's response to mean warming would mimic its response to  
604 interannual temperature anomalies. Models that are most responsive to interannual temperature  
605 variability would therefore be expected to also be the most sensitive to mean temperature  
606 changes.

607  
608 For example, consider two models: Model A (which simulates higher yields in warm years and  
609 thus whose response is positively correlated with interannual temperatures) and Model B (which  
610 simulates lower yields in warm years and thus whose response is negatively correlated with  
611 interannual temperatures). A temporal proxy assumption would anticipate that Model A would

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

612 have more positive simulated yield changes (as a percentage of the historical simulations' yields)  
613 than Model B if both were exposed to warmer mean conditions. Likewise, if both models were  
614 simulated under cooler mean conditions Model A would have more negative yield changes than  
615 Model B. These comparisons between climate variability sensitivities and climate change  
616 responses are informative not only for the relationship of a single given model, but the pattern of  
617 the full ensemble provides a basis on which to evaluate model consistency and simple statistical  
618 modeling approaches.

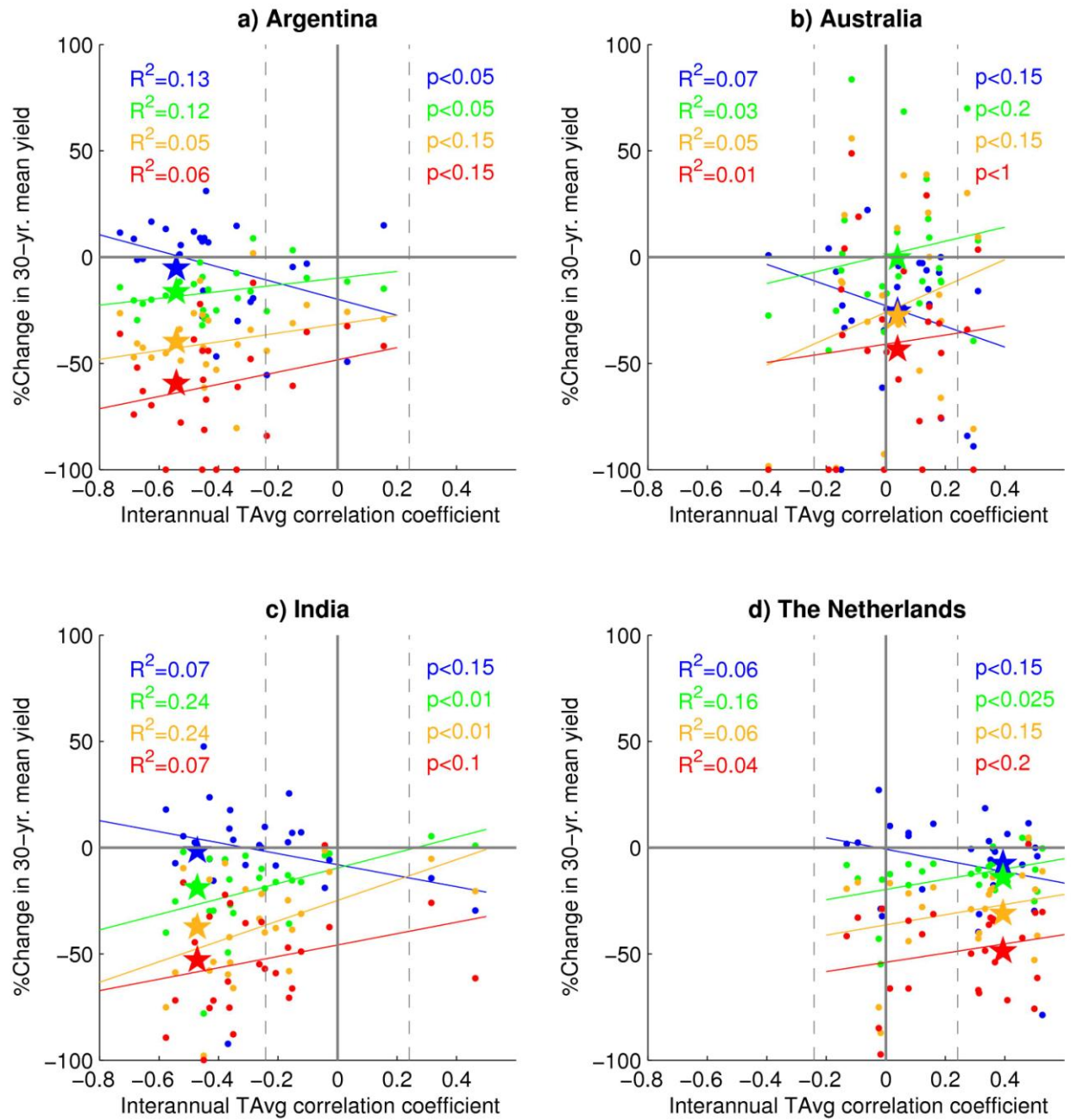
619  
620 The 27 wheat models' interannual temperature sensitivity and mean temperature change  
621 responses are compared for each of the temperature sensitivity tests and each of the four  
622 locations in **Figure 7**, with each dot representing a single wheat model. A model's position on  
623 the x-axis represents the correlation of its interannual yields against growing season temperature  
624 anomalies in the 1980-2010 period, and its position on the y-axis represents the percentage  
625 change in mean yield (over the 30 growing seasons) for each of the temperature sensitivity tests  
626 in comparison to the 1980-2010 mean yield (with CO<sub>2</sub> held at historical concentrations of 360  
627 ppm). A linear fit is also drawn for each color-coded sensitivity test (quadratic fits were not  
628 substantially better).

629  
630 As expected, the slopes of the linear fits indicate that models with greater interannual  
631 temperature sensitivity are more sensitive to mean temperature changes. The +3°C, +6°C, and  
632 +9°C sensitivity tests' linear fits have a positive slope at all sites. This indicates that the mean  
633 warming tended to lead to relatively higher simulated grain yields in models with more positive  
634 correlations between interannual temperature and grain yield compared to models with more

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

635 negative correlations (which had lower simulated grain yield in the sensitivity tests). Also as  
636 expected, the -3°C sensitivity test’s linear fit has a negative slope, as decreases in mean  
637 temperature lead to larger grain yield losses when models’ interannual temperature anomalies are  
638 more positively correlated with yields compared to models.

639  
640



**Figure 7:** Comparison between each model's Pearson's correlation coefficient of interannual temperature and grain yield with its response to mean temperature change sensitivity tests (-3°C, blue; +3°C, green; +6°C, orange; +9°C, red; each compared against 1981-2010 historical period at single location in each country). Ensemble averages for each sensitivity test are represented by a star, and colored lines represent the least-squares linear fit for each sensitivity test with  $R^2$  correlation and t-test significance documented for each fit. Vertical dashed lines indicate t-test significance at the 90<sup>th</sup>-percentile level for interannual correlations between average temperature and simulated

1  
2  
3  
4 648 grain yield. The  $R^2$  correlation and significance level for the fitted slope of each least-squares-fitted line is also  
5  
6 649 presented in text of the corresponding color in each panel. p-levels presented for the slope were the lowest possible  
7  
8 650 among 0.005, 0.01, 0.025, 0.05, 0.1, 0.15, 0.2, 0.25, and 1.  
9

10 651  
11  
12 652 While the slopes of these lines support the use of temporal proxies for climate impact analyses,  
13  
14  
15 653 other aspects of the analysis cast serious doubt on the utility of the temporal proxy approach  
16  
17  
18 654 (even when  $CO_2$  is held constant). Firstly, there is a dramatic spread among the 27 wheat models  
19  
20 655 around the fitted line, with the sign of many models' mean temperature change responses  
21  
22  
23 656 opposite from what would be predicted by the interannual temperature response. As shown in  
24  
25 657 Figure 7,  $R^2$  correlations are quite low (between 0 and 0.24), with lowest values in the +9°C  
26  
27  
28 658 sensitivity test. Correlations are particularly low in Australia ( $R^2 \leq 0.07$ ) where interannual  
29  
30 659 temperature sensitivity was weak in most models, and are highest in the +3°C and +6°C  
31  
32 660 sensitivity tests for India ( $R^2 = 0.24$ ) where irrigation likely enabled a stronger temperature  
33  
34  
35 661 signal. *t*-test evaluations of the least-squares fit reveal many instances where the slopes are not  
36  
37 662 statistically significant at the  $p=0.05$  level, particularly in Australia and for the higher  
38  
39  
40 663 temperature change sensitivity tests (where only India is significant at the  $p<0.1$  level).  
41  
42 664 Together, these low correlations and the weak significance of fitted slopes suggest that the  
43  
44  
45 665 temporal proxy cannot be reliably applied, especially for conditions that are substantially warmer  
46  
47 666 than the calibration period.

48  
49 667  
50  
51 668 Secondly, a temporal proxy would predict that models with no sensitivity to interannual  
52  
53  
54 669 temperature variability would have no response to climate change (as represented by the  
55  
56  
57 670 temperature sensitivity tests), and therefore all linear fits should intersect at the origin of the  
58  
59 671 axes. This is not the case as nearly all temperature sensitivity test lines fall below the origin with  
60  
61

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

672 increasing distance as temperatures rise, suggesting that additional factors impart a mean grain  
673 yield reduction above what would be expected from examining the impacts of historical  
674 temperature variability. Several potential explanations for these differences merit further study.

675  
676 A first candidate factor is that this simple temporal proxy based solely on temperature lends itself  
677 to biases as a result of interdependence of climate variables (Sheehy et al., 2006). For example,  
678 temperature anomalies may correlate with yield losses only because they coincide with dry  
679 seasons, which would suggest that a rainfall-based empirical model would be more appropriate.  
680 Interdependence of climate variables would somewhat explain the deviations of the wheat  
681 models around the least-squares fitted lines in Figure 7 as the interannual correlation would not  
682 be solely a temperature sensitivity. This factor cannot explain the extent of these deviations,  
683 however, nor is this explanation sufficient to explain the offset at the origin.

684  
685 A second factor is the non-linearity in grain yield responses as mean climate change pushes  
686 systems beyond critical thresholds and tipping points, some of which may not have been present  
687 in the historical conditions. Within each temperature sensitivity test there are 30 years of  
688 climate variability including warm seasons with extreme events that are amplified by an  
689 increasing mean temperature and which may have a disproportionate impact on the mean yield  
690 shift. In combination, the mean warming and interannual extremes can produce conditions  
691 never experienced during the 1981-2010 period. In many cases this leads to a non-linear impact  
692 on grain yields beyond a simple extrapolation of interannual proxies (Porter and Semenov,  
693 1995). For example, Lobell et al. (2012) found an acceleration of leaf senescence in Indian  
694 wheat during extreme heat events beyond what would have been expected from average

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

695 temperatures alone. Interactions with other variables can also compound yield losses. Chief  
696 among these are increases in water stress during critical growth stages, as warmer temperatures  
697 lead to increased vapor pressure deficit and higher potential evapotranspiration (although  
698 accumulated water requirements may be partially counter-balanced by a shorter growing  
699 season). Non-linear effects could be identified if particular years in the sensitivity tests  
700 experienced much larger losses than the average year (compared to the historical climate).  
701 Thresholds and plant stresses at critical growth stages can also lead to complete loss of grain  
702 yields, as is clear in the number of models reporting 100% grain yield loss under the highest  
703 temperature conditions (Figure 7).

704

705 A third factor relates to different responses of grain yield to temperature variability and change  
706 during different parts of the crop growing season or during different parts of the year. This is  
707 probably particularly relevant for crops with a long growing period such as winter wheat in the  
708 Netherlands. An example of this is winter wheat in Denmark, where Kristensen et al. (2011)  
709 found a positive response of yield to increased temperature at low temperatures during winter,  
710 but a highly negative response during summer. Also Liu et al. (2013) found differential effects  
711 of warming on winter wheat yield in the North China Plain depending on whether the warming  
712 mainly affected winter or summer conditions. The effects of warming for crops that have long  
713 growing seasons with large seasonal differences may therefore be obscured by positive effects  
714 of warming in some parts of the growing season and negative ones in other parts of the growth  
715 period.

716

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

717 A final candidate factor for the differences between interannual temperature variability and mean  
718 warming is the extent of within-season climate variability. In the historical record extremely  
719 warm seasons tend to be only marginally warm on the average day but feature a substantial heat  
720 wave (or several), which has a fundamentally different effect on plant function from that of a  
721 season where a slight warming is relentless (even if the average temperature is the same). With  
722 prolonged warming maturation is accelerated and yields may be reduced as a result of lower net  
723 radiation interception. There is also an increased chance that warm temperatures will negatively  
724 affect key phenological stages and/or interact with precipitation or solar radiation to create  
725 evaporative demand that the plants cannot meet. These alterations to phenological development  
726 and/or heat and water stresses can have cascading effects on plant growth throughout the season  
727 with net yield reductions on average compared to the historical temperature variability. The  
728 models respond to high temperatures according to a large variety of parameterizations (Alderman  
729 et al., 2013), with responses to extreme heat an area in particular need of development (Lobell et  
730 al., 2012).

731

**4. Conclusions and next steps**

732 Analysis of the 27 models participating in the AgMIP Wheat Model Intercomparison Pilot  
733 reveals substantial differences in the ways that models respond to interannual variations in  
734 rainfall, temperature, and solar radiation at four diverse locations. These differences provide  
735 useful context to differences in the abilities of the same models to reproduce detailed field  
736 observations (Martre et al., 2015) and climate change responses (Asseng et al., 2013, 2015). The  
737 large differences apparent in interannual climate sensitivity suggest that multiple years of  
738 consistent field trials are desirable to enable proper initialization of field conditions, and field



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

740 experiments during extreme conditions would benefit the calibration of crop models for both  
741 mean yields and interannual variability. Such long-term agricultural research datasets are rare,  
742 unfortunately, so in typical applications such as those done here it is likely that any biases in  
743 calibration are amplified when a single-year's calibration is used for multiple seasons. It is  
744 therefore useful to take advantage of the tendency of multi-model ensemble statistics to reduce  
745 overall errors beyond the calibration period.

746  
747 The AgMIP Wheat Pilot offers a far larger multi-model sample than would be expected in the  
748 applications for which each of the participating models was designed; however several of the  
749 interannual response results help guide the formation of practical subsets and application  
750 protocols. Although calibration information has been shown to reduce errors in mean yields and  
751 details in crop growth (Asseng et al., 2013), the results presented here suggest that interannual  
752 yield variability for most models is not strongly affected by the availability of more detailed field  
753 observations (e.g., evapotranspiration, biomass, leaf-area index, plant available soil moisture) for  
754 calibration. This is encouraging as high-information field trials are much less common. Adding  
755 a second (and third) wheat model dramatically increases the likelihood that the simulated results  
756 will reproduce the interannual behavior of the full 27-model ensemble, with a diminishing  
757 benefit to efforts that utilize additional models beyond that. This information is directly relevant  
758 to the design of new studies looking to take advantage of multi-model ensemble statistics despite  
759 resource constraints, including AgMIP efforts to form crop modeling tools that may link with  
760 global agricultural monitoring and outlooks on a sub-seasonal to seasonal scale (Singh et al.,  
761 2012; Vitart et al., 2012). Use of an ensemble also highlights the sensitivity of simulated yields

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

762 to interannual climate variability as common features rise above the ensemble’s diminished noise  
763 more easily than the individual models’ larger noise.

764  
765 The wheat models demonstrate several common patterns of climate variability response at each  
766 tested location. In some cases there is a fundamental disagreement between models about  
767 whether grain yield responds positively or negatively to a given anomaly, although  
768 interdependence of climate variables (e.g., wet and cool years vs. hot and dry years) muddles the  
769 picture. Even when two models respond in a very similar manner at one location, differences in  
770 calibration method and quality, parameters, model structure, and environmental conditions can  
771 lead to strong deviations in model response at other sites. These results therefore suggest that  
772 there are still strong differences in wheat models’ climate sensitivities, and that further work is  
773 needed to create models that are truly applicable across a wide range of current and future  
774 conditions. The analysis presented here focuses on mean growing season climate anomalies at  
775 four locations; however consideration of intra-seasonal variability and extremes (e.g., heat  
776 waves, dry spells, frosts, floods, waterlogging, monsoon dynamics) require further study.  
777 Comparing multi-model simulation experiments against long-term field trials (e.g., Dobermann  
778 et al., 2000) would also be desirable in order to provide true observations upon which to evaluate  
779 simulated outputs (rather than assuming the value of the ensemble average as done here).

780  
781 The effects of interannual temperature variability and mean climate warming were shown to be  
782 only weakly related among the 27 wheat models, indicating that a temporal proxy for climate  
783 change is likely oversimplified. State-of-the-art empirical models use far more than interannual  
784 temperature for climate impacts projection, however these findings underscore the importance of

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

785 considering complex interactions between variables and non-linear responses that may not be  
786 present in the historical period datasets to which models are fit. Further work is needed to  
787 elucidate additional physiological factors that differentiate the effects of a warm season from  
788 those of a warmer climate (Porter and Semenov et al., 2005).

789  
790 Follow-on phases of the AgMIP Wheat Pilot are focusing on more sites and experiments  
791 designed to better distinguish between heat waves and warmer mean climate conditions. The  
792 analyses presented here would also be of interest for other completed AgMIP Crop Model Pilots  
793 (e.g., for maize, Bassu et al., 2014; rice, Li et al., 2014; and sugarcane, Singels et al., 2013) as  
794 well as pilots planned for millet and sorghum, potato, canola, and grasslands. AgMIP's  
795 Coordinated Climate-Crop Modeling Project (C3MP; Ruane et al., 2014; McDermid et al., 2015)  
796 and Global Gridded Crop Model Intercomparison (GGCMI; Rosenzweig et al., 2014; Elliott et  
797 al., 2015), as well as the impact response surface studies conducted in FACCE MACSUR  
798 (Pirttioja et al., 2015) provide additional fora in which to compare climate sensitivities across  
799 multiple locations and crop models, assuming that observational yield data also are available for  
800 those points or aggregated grid cells. This study's yield response analyses are currently being  
801 applied to GGCMI's historical period intercomparison, helping to determine the causes for  
802 differences in interannual yield variation for more than a dozen models with global coverage of  
803 multiple crops (Elliott et al., 2015). Wheat model development would benefit from a future  
804 intercomparison centered upon a region where long-term variety trials overlap with similar  
805 detailed field experiments so that calibration and the response to interannual climate variability  
806 may be more comprehensively evaluated. Of particular interest would be the way in which

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

807 interannual yield observations affect calibration and the resulting climate variability and climate  
808 change sensitivities.

809  
810 Results from this study underscore the need for model intercomparison results to avoid  
811 anonymity in order to enable careful analysis of structural and parameter differences that cause  
812 differences in yield response. Current and future phases of the AgMIP Wheat intercomparisons  
813 no longer hold the models anonymous, and evaluation of the mechanisms driving different  
814 climate responses is a crucial line of continuing inquiry (as was performed for the AgMIP Rice  
815 Pilot; Li et al., 2015). Through these activities the efforts of the AgMIP Wheat Pilot will better  
816 accomplish integrated assessments of climate impact on the agricultural sector.

817  
818 **Acknowledgements**

819 We acknowledge the efforts of AgMIP Leaders Cynthia Rosenzweig, Jim Jones, John Antle, and  
820 Jerry Hatfield in their efforts to initiate the AgMIP Wheat Pilot and encourage explorations such  
821 as these. AgMIP coordination and climate activities were funded in part by a grant from the UK  
822 Department for International Development. Support for many European participants in the  
823 AgMIP Wheat Pilot was provided by the Modelling European Agriculture with Climate Change  
824 for Food Security (MACSUR) knowledge hub within the Joint Research Programming Initiative  
825 on Agriculture, Food Security and Climate Change (FACCE-JPI). Alex Ruane's contributions  
826 were supported by the NASA Earth Sciences Research Program (#281945.02.03.03.96) and the  
827 NASA Modeling, Analysis, and Prediction Program (#509496.02.08.04.24). Participation of  
828 Claudio Stöckle and CropSyst simulations were supported by the project Regional Approaches to

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

829 Climate Change for Pacific Northwest Agriculture (REACCH-PNA), funded through award

830 #2011-68002-30191 from the US National Institute for Food and Agriculture.

831

832

1  
2  
3  
4 **833** **References**

- 5  
6  
7 834 Aggarwal, P., B. Banerjee, M. Daryaei, A. Bhatia, A. Bala, S. Rani, S. Chander, H. Pathak, and  
8  
9 835 N. Kalra, 2006: InfoCrop: A dynamic simulation model for the assessment of crop yields,  
10  
11 836 losses due to pests, and environmental impact of agro-ecosystems in tropical  
12  
13  
14 837 environments. II. Performance of the model. *Agricultural Systems* 89: 47-67. doi:  
15  
16 838 10.1016/j.agsy.2005.08.003|10.1016/j.agsy.2005.08.003.  
17  
18  
19 839 Alderman, P.D., E. Quilligan, S. Asseng, F. Ewert, and M.P. Reynolds, 2013: Proceedings of the  
20  
21 840 Workshop: Modeling Wheat Response to High Temperature. CIMMYT, El Batan,  
22  
23 841 Texcoco, Mexico. 19-21 June, 2013. 138pp.  
24  
25  
26 842 Angulo C., R. Rötter, R. Lock, A. Enders, S. Fronzek, and F. Ewert, 2013: Implication of crop  
27  
28 843 model calibration strategies for assessing regional impacts of climate change in Europe.  
29  
30  
31 844 *Agricultural and Forest Meteorology* **170**, 32-46.  
32  
33  
34 845 Arnold, G.W., de Wit, C.T., 1976. Critical evaluation of system analysis in ecosystem  
35  
36 846 management. Simulation Monographs, Wageningen, The Netherlands.  
37  
38 847 Asseng, S., H. van Keulen, and W. Stol, 2000: Performance and application of the APSIM  
39  
40  
41 848 Nwheat model in the Netherlands. *European Journal of Agronomy* 12: 37-54.  
42  
43 849 Asseng, S., P.D. Jamieson, B. Kimball, P. Pinter, K. Sayre, J.W. Bowden, and S.M. Howden,  
44  
45 850 2004: Simulated wheat growth affected by rising temperature, increased water deficit and  
46  
47  
48 851 elevated atmospheric CO<sub>2</sub>. *Field Crops Research* 85: 85-102.  
49  
50  
51 852 Asseng, S., B.A. Keating, I.R.P. Fillery, P.J. Gregory, J.W. Bowden, N.C. Turner, J.A. Palta, and  
52  
53 853 D.G. Abrecht, 1998: Performance of the APSIM-wheat model in Western Australia. *Field*  
54  
55 854 *Crops Research* 57: 163-179.  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

855 Asseng, S., H. van Keulen, and W. Stol, 2000: Performance and application of the APSIM  
856 Nwheat model in the Netherlands. *European Journal of Agronomy* 12: 37-54.

857 Asseng, S., F. Ewert, C. Rosenzweig, J.W. Jones, J.L. Hatfield, A.C. Ruane, K.J. Boote, P.  
858 Thorburn, R.P. Rötter, D. Cammarano, N. Brisson, B. Basso, P. Martre, P.K. Aggarwal,  
859 C. Angulo, P. Bertuzzi, C. Biernath, A.J. Challinor, J. Doltra, S. Gayler, R. Goldberg, R.  
860 Grant, L. Heng, J. Hooker, L.A. Hunt, J. Ingwersen, R.C. Izaurralde, K.C. Kersebaum, C.  
861 Müller, S. Naresh Kumar, C. Nendel, G. O’Leary, J.E. Olesen, T. M. Osborne, T.  
862 Palosuo, E. Priesack, D. Ripoche, M.A. Semenov, I. Shcherbak, P. Steduto, C. Stöckle, P.  
863 Stratonovitch, T. Streck, I. Supit, F. Tao, M. Travasso, K. Waha, D. Wallach, J.W. White,  
864 J.R. Williams and J. Wolf, 2013: Uncertainty in simulating wheat yields under climate  
865 change. *Nature Clim. Change* 3, 827-832. doi: 10.1038/NCLIMATE1916

866 Asseng, S., F. Ewert, P. Martre, R.P. Rötter, D.B. Lobell, D. Cammarano, B.A. Kimball, M.J.  
867 Ottman, G.W. Wall, J.W. White, M.P. Reynolds, P.D. Alderman, P.V. V. Prasad, P.K.  
868 Aggarwal, J. Anothai, B. Basso, C. Biernath, A.J. Challinor, G. De Sanctis, J. Doltra, E.  
869 Fereres, M. Garcia-Vila, S. Gayler, G. Hoogenboom, L.A. Hunt, R.C. Izaurralde, M.  
870 Jabloun, C.D. Jones, K.C. Kersebaum, A-K. Koehler, C. Müller, S. Naresh Kumar, C.  
871 Nendel, G. O’Leary, J.E. Olesen, T. Palosuo, E. Priesack, E. Eyshi Rezaei, A.C. Ruane,  
872 M.A. Semenov, I. Shcherbak, C. Stockle, P. Stratonovitch, T. Streck, I. Supit, F. Tao, P.J.  
873 Thorburn, K. Waha, E. Wang, D. Wallach, J. Wolf, Z. Zhao, and Y. Zhu, 2015: Rising  
874 temperatures reduce global wheat production. *Nature Clim. Change*, 5(2), 143-147,  
875 doi:10.1038/nclimate2470.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

876 Basso B, Cammarano D, Troccoli A, Chen D, Ritchie J (2010) Long-term wheat response to  
877 nitrogen in a rainfed Mediterranean environment: Field data and simulation analysis.  
878 European Journal of Agronomy 33: 132-138. doi: DOI 10.1016/j.eja.2010.04.004.

879 Bassu S., N. Brisson, J-L. Durand, K.J. Boote, J. Lizaso, J.W. Jones, C. Rosenzweig, A.C.  
880 Ruane, M. Adam, C. Baron, B. Basso, C. Biernath, H. Boogaard, S. Conijn, M. Corbeels,  
881 D. Deryng, G. De Sanctis, S. Gayler, P. Grassini, J. Hatfield, S. Hoek, C. Izaurrealde, R.  
882 Jongschaap, A. Kemanian, K.C. Kersebaum, D. Makowski, C. Müller, C. Nendel, E.  
883 Priesack, M.V. Pravia, F. Sau, I. Shcherbak, F. Tao, E. Teixeira, D. Timlin, and K. Waha,  
884 2014: Do various maize crop models give the same responses to climate change factors?  
885 *Global Change Biology* **20**, 2301–2320, doi: 10.1111/gcb.12520

886 Berntsen, J., B. Petersen, B. Jacobsen, J. Olesen, and N. Hutchings, 2003: Evaluating nitrogen  
887 taxation scenarios using the dynamic whole farm simulation model FASSET.  
888 *Agricultural Systems* 76: 817-839. doi: 10.1016/S0308-521X(02)00111-  
889 7|10.1016/S0308-521X(02)00111-7.

890 Biernath, C., S. Gayler, S. Bittner, C. Klein, P. Hogy, A. Fangmeier, and E. Priesack, 2011:  
891 Evaluating the ability of four crop models to predict different environmental impacts on  
892 spring wheat grown in open-top chambers. *European Journal of Agronomy* 35: 71-82.  
893 doi: DOI 10.1016/j.eja.2011.04.001.

894 Bondeau, A., P.C. Smith, S. Zaehle, S. Schaphoff, W. Lucht, W. Cramer, and D. Gerten, 2007:  
895 Modelling the role of agriculture for the 20th century global terrestrial carbon balance.  
896 *Global Change Biology* 13: 679-706.



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

897 Boogaard, H.L., C.A. Van Diepen, R.P. Rötter, J.C.M.A. Cabrera, and H.H. Van Laar (eds),  
898 1998: User's guide for the WOFOST 7.1 crop growth simulation model and WOFOST  
899 control center 1.5. Winand Staring Centre, Wageningen, The Netherlands, 144pp.

900 Boote, K.J., C.H. Porter, J.W. Jones, P.J. Thorburn, K.C. Kersebaum, G. Hoogenboom, J.W.  
901 White, and J.L. Hatfield, 2015: Sentinel Site Data for Model Improvement – Definition  
902 and Characterization. In: Hatfield, J.L., Fleisher, D. (Eds.), Improving Modeling Tools to  
903 Assess Climate Change Effects on Crop Response, Advances in Agricultural Systems  
904 Modeling. ASA, CSSA, and SSSA, Madison, WI, USA.  
905 doi:10.2134/advagricysystem7.2014.0019

906 Brisson, N., C. Gary, E. Justes, R. Roche, B. Mary, D. Ripoche, D. Zimmer, J. Sierra, P.  
907 Bertuzzi, P. Burger, F. Bussiere, Y. Cabidoche, P. Cellier, P. Debaeke, J. Gaudillere, C.  
908 Henault, F. Maraux, B. Seguin, and H. Sinoquet, 2003: An overview of the crop model  
909 STICS. European Journal of Agronomy 18: 309-332. doi: PII S1161-0301(02)00110-7.

910 Brisson, N., B. Mary, D. Ripoche, M. Jeuffroy, F. Ruget, B. Nicoullaud, P. Gate, F. Devienne-  
911 Barret, R. Antonioletti, C. Durr, G. Richard, N. Beaudoin, S. Recous, X. Tayot, D.  
912 Plenet, P. Cellier, J. Machet, J. Meynard, and R. Delecolle, 1998: STICS: a generic  
913 model for the simulation of crops and their water and nitrogen balances. I. Theory and  
914 parameterization applied to wheat and corn. Agronomie 18: 311-346.

915 Challinor, A., T. Wheeler, P. Craufurd, J. Slingo, and D. Grimes, 2004: Design and optimisation  
916 of a large-area process-based model for annual crops. Agricultural and Forest  
917 Meteorology 124: 99-120. doi: DOI 10.1016/j.agrformet.2004.01.002.

918 Challinor, A.J., and Coauthors, 2014a: A meta-analysis of crop yield under climate change and  
919 adaptation. Nature Climate Change 4(4), 287-291. doi:10.1038/NCLIMATE2153

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

920 Challinor, A.J., P. Martre, S. Asseng, P. Thornton, and F. Ewert, 2014b: Making the most of  
921 climate impact ensembles. *Nature Climate Change* 4(2), 77-80.

922 Dobermann, A., D. Dawe, R.P. Rötter, and K.G. Cassman, 2000: Reversal of rice yield decline  
923 in a long-term continuous cropping experiment. *Agronomy Journal*, 92, 633-643.

924 Donatelli, M., and Co-authors, 2010: A Component-Based Framework for Simulating  
925 Agricultural Production and Externalities. *Environmental and Agricultural Modelling:  
926 Integrated Approaches For Policy Impact Assessment*: 63-108. doi: 10.1007/978-90-481-  
927 3619-3\_4|10.1007/978-90-481-3619-3\_4.

928 Elliott, J., C. Müller, D. Deryng, J. Chryssanthacopoulos, K.J. Boote, M. Büchner, I. Foster, M.  
929 Glotter, J. Heinke, T. Iizumi, R.C. Izaurralde, N.D. Mueller, D.K. Ray, C. Rosenzweig,  
930 A.C. Ruane, and J. Sheffield, 2015: The Global Gridded Crop Model Intercomparison:  
931 Data and modeling protocols for Phase 1 (v1.0). *Geosci. Model Dev.*, 8, 261-277,  
932 doi:10.5194/gmd-8-261-2015.

933 Evans, L.T., and R.A. Fischer, 1999: Yield potential: Its definition, measurement, and  
934 significance. *Crop Science*, 39, 1544-1551.

935 Ewert, F., C. Angulo, C. Rumbaur, R. Lock, A. Enders, M. Andenäuer, T. Heckeley, M. van  
936 Ittersum, J. Wolf, R. Rötter, 2011a: Assessing the adaptive capacity of agriculture in the  
937 Netherlands to the impacts of climate change under different market and policy scenarios.  
938 AgriAdapt Project Reports no 2 & 3 Dutch Research Program Climate Change and  
939 Spatial Planning: 49.

940 Ewert F., M.K. van Ittersum, T. Heckeley, O. Therond, I. Bezlepkina, and E. Andersen, 2011b:  
941 Scale changes and model linking methods for integrated assessment of agri-  
942 environmental systems. *Agriculture, Ecosystems and Environment* **142**, 6-17.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

943 Ewert, F., R.P. Rötter, M. Bindi, H. Webber, M. Trnka, K-C. Kersebaum, J.E. Olesen, M.K. van  
944 Ittersum, S. Janssen, M. Rivington, M. Semenov, D. Wallach, J.R. Porter, D. Stewart, J.  
945 Verhagen, T. Gaiser, T. Palosuo, F. Tao, C. Nendel, P-P. Roggero, L. Bartosova, and S.  
946 Asseng, 2015: Crop modelling for integrated assessment of risk to food production from  
947 climate change. *Environmental Modelling & Software*, 72, 287-303.  
948 doi:10.1016/j.envsoft.2014.12.003

949 Fader, M., S. Rost, C. Muller, A. Bondeau, and D. Gerten, 2010: Virtual water content of  
950 temperate cereals and maize: Present and potential future patterns. *Journal of Hydrology*  
951 384: 218-231. doi: 10.1016/j.jhydro1.2009.12.011.

952 Ferrise, R., A. Triossi, P. Stratonovitch, M. Bindi, and P. Martre. 2010: Sowing date and nitrogen  
953 fertilisation effects on dry matter and nitrogen dynamics for durum wheat: An  
954 experimental and simulation study. *Field Crops Res* 117: 245-257.

955 Goudriaan, J., H.H. Van Laar (eds), 1994: *Modelling Potential Crop Growth Processes.*  
956 *Textbook With Exercises.* Kluwer Academic Publishers, Dordrecht, The Netherlands. pp.  
957 238.

958 Grant, R., B. Kimball, M. Conley, J. White, G. Wall, and M. Ottman, 2011: Controlled Warming  
959 Effects on Wheat Growth and Yield: Field Measurements and Modeling. *Agronomy*  
960 *Journal* 103: 1742-1754. doi: 10.2134/agronj2011.0158.

961 He, J., J. Le Gouis, P. Stratonovitch, V. Allard, O. Gaju, E. Heumez, S. Orford, S. Griffiths, J.W.  
962 Snape, M.J. Foulkes, M.A. Semenov, and P. Martre, 2012: Simulation of environmental  
963 and genotypic variations of final leaf number and anthesis date for wheat. *European*  
964 *Journal of Agronomy* 42, 22-33.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

965 He, J., P. Stratonovitch, V. Allard, M.A. Semenov, and P. Martre, 2010: Global Sensitivity  
966 Analysis of the Process-Based Wheat Simulation Model SiriusQuality1 Identifies Key  
967 Genotypic Parameters and Unravels Parameters Interactions. *Procedia - Social and  
968 Behavioral Sciences* 2: 7676-7677.

969 Hoogenboom, G., and J. White, 2003: Improving physiological assumptions of simulation  
970 models by using gene-based approaches. *Agronomy Journal* 95: 82-89.

971 Hunt, L.A., and S. Pararajasingham, 1995: CROPSIM-wheat - a model describing the growth  
972 and development of wheat. *Canadian Journal of Plant Science* 75: 619-632.

973 Jamieson, P., and M. Semenov, 2000: Modelling nitrogen uptake and redistribution in wheat.  
974 *Field Crops Research* 68: 21-29. doi: 10.1016/S0378-4290(00)00103-9.

975 Jamieson, P., M. Semenov, I. Brooking, and G. Francis, 1998: Sirius: a mechanistic model of  
976 wheat response to environmental variation. *European Journal of Agronomy* 8: 161-179.  
977 doi: 10.1016/S1161-0301(98)00020-3.

978 Jones, J., G. Hoogenboom, C. Porter, K. Boote, W. Batchelor, L. Hunt, P. Wilkens, U. Singh, A.  
979 Gijsman, and J. Ritchie, 2003: The DSSAT cropping system model. *European Journal of  
980 Agronomy* 18: 235-265. doi: PII S1161-0301(02)00107-7.

981 Keating, B.A., and Co-Authors, 2003: An overview of APSIM, a model designed for farming  
982 systems simulation. *European Journal of Agronomy* 18: 267-288.

983 Kersebaum, K., 1995: Application of a simple management model to simulate water and nitrogen  
984 dynamics. *Ecological Modelling* 81: 145-156.

985 Kersebaum, K., 2007: Modelling nitrogen dynamics in soil-crop systems with HERMES.  
986 *Nutrient Cycling in Agroecosystems* 77: 39-52. doi: DOI 10.1007/s10705-006-9044-8.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

987 Kersebaum, K.C., 2011: Special features of the HERMES model and additional procedures for  
988 parameterization, calibration, validation, and applications. Ahuja, LR and Ma, L (eds)  
989 Methods of introducing system models into agricultural research Advances in  
990 Agricultural Systems Modeling Series 2, Madison (ASA-CSSA-SSSA): 65-94.

991 Kersebaum, K.C., and A.J. Beblik, 2001: Performance of nitrogen dynamics model applied to  
992 evaluate agricultural management practices. Lewis Publishers.

993 Kersebaum, K.C., K.J. Boote, J.S. Jorgenson, C. Nendel, M. Bindi, C. Frühauf, T. Gaiser, G.  
994 Hoogenboom, C. Kollas, J.E. Olesen, R.P. Rötter, F. Ruget, P. Thorburn, M. Trnka, M.  
995 Wegehenkel, 2015: Analysis and classification of data sets for calibration and validation  
996 of agro-ecosystem models. *Environmental Modelling & Software*, 72, 402-417. DOI:  
997 10.1016/j.envsoft.2015.05.009.

998 Kiniry, J., D. Major, R. Izaurrealde, J. Williams, P. Gassman, M. Morrison, R. Bergentine, and R.  
999 Zentner, 1995: EPIC model parameters for cereal, oilseed, and forage crops in the  
1000 northern great-plains region. *Canadian Journal of Plant Science* 75: 679-688.

1001 Kollas, C., K.-C. Kersebaum, C. Nendel, K. Manevski, C. Müller, T. Palosuo, C.M. Armas-  
1002 Herrera, N. Beaudoin, M. Bindi, M. Charfeddine, T. Conradt, J. Constantin, J. Eitzinger,  
1003 F. Ewert, R. Ferrise, T. Gaiser, I. Garcia de Cortazar-Atauri, L. Giglio, P. Hlavinka, H.  
1004 Hoffmann, M.P. Hoffmann, M. Launay, R. Manderscheid, B. Mary, W. Mirschel, M.  
1005 Moriondo, J.E. Olesen, I. Öztürk, A. Pacholski, D. Ripoche-Wachter, P.P. Roggero, S.  
1006 Roncossek, R.P. Rötter, F. Ruget, B. Sharif, M. Trnka, D. Ventrella, K. Waha, M.  
1007 Wegehenkel, H-J. Weigel, and L. Wu, 2015: Crop rotation modelling – a European  
1008 model intercomparison. *Eur. J. Agron.* 70, 98-111. DOI: 10.1016/j.eja.2015.06.007

1  
2  
3  
4 1009 Kristensen, K., K. Schelde, and J.E. Olesen, 2011: Winter wheat yield response to climate  
5  
6  
7 1010 variability in Denmark. *Journal of Agricultural Science* 149: 33-47.  
8  
9 1011 Latta, J., and G. O'Leary, 2003: Long-term comparison of rotation and fallow tillage systems of  
10  
11  
12 1012 wheat in Australia. *Field Crops Research* 83: 173-190. doi: 10.1016/S0378-  
13  
14 1013 4290(03)00073-X|10.1016/S0378-4290(03)00073-X.  
15  
16 1014 Lawless, C., M. Semenov, and P. Jamieson, 2005: A wheat canopy model linking leaf area and  
17  
18  
19 1015 phenology. *European Journal of Agronomy* 22: 19-32. doi: 10.1016/j.eja.2003.11.004.  
20  
21 1016 Li, S., T. Wheeler, A. Challinor, E. Lin, Y. Xu, and H. Ju, 2010: Simulating the Impacts of  
22  
23  
24 1017 Global Warming on Wheat in China Using a Large Area Crop Model. *Acta*  
25  
26 1018 *Meteorologica Sinica* 24: 123-135.  
27  
28  
29 1019 Li, T., T. Hasegawa, X. Yin, Y. Zhu, K. Boote, M. Adam, S. Bregaglio, S. Buis, R. Confalonieri,  
30  
31 1020 T. Fumoto, D. Gaydon, M. Marcaida, III, H. Nakagawa, P. Oriol, A.C. Ruane, F. Ruget,  
32  
33  
34 1021 B. Singh, U. Singh, L. Tang, F. Tao, P. Wilkens, H. Yoshida, Z. Zhang, and B. Bouman,  
35  
36 1022 2015: Uncertainties in predicting rice yield by current crop models under a wide range of  
37  
38 1023 climatic conditions. *Glob. Change Biol.*, 21, no. 3, 1328-1341, doi:10.1111/gcb.12758.  
39  
40  
41 1024 Liu, L., C. Hu, J.E. Olesen, Z. Ju, P. Yang, Y. Zhang, 2013. Warming and nitrogen fertilization  
42  
43 1025 effects on winter wheat yields in northern China varied between four years. *Field Crops*  
44  
45 1026 *Research* 151: 56-64.  
46  
47  
48 1027 Lobell, D.B., and M.B. Burke, 2010: On the use of statistical models to predict crop yield  
49  
50  
51 1028 responses to climate change. *Agricultural and Forest Meteorology* 150, 1443–1452.  
52  
53 1029 Lobell, D.B., A. Sibley, and J.I. Ortiz-Monasterio, 2012: Extreme heat effects on wheat  
54  
55 1030 senescence in India. *Nature Climate Change* 2, 186-189. doi:10.1038/NCLIMATE1356  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 1031 Martre, P., P.D. Jamieson, M.A. Semenov, R.F. Zyskowski, J.R. Porter, and E. Triboi, 2006:  
5  
6  
7 1032 Modelling protein content and composition in relation to crop nitrogen dynamics for  
8  
9 1033 wheat. *Eur J Agron* 25: 138-154.  
10  
11 1034 Martre, P., and Coauthors, 2014: Error of multimodel ensembles of wheat growth: more models  
12  
13 are better than one. *Global Change Biology*, 21, 911–925, doi: 10.1111/gcb.12768.  
14 1035  
15  
16 1036 McDermid, S.P., A.C. Ruane , C. Rosenzweig, N.I. Hudson, M.D. Morales, P. Agalawatte, S.  
17  
18  
19 1037 Ahmad, L.R. Ahuja, I. Amien, S.S. Anapalli, J. Anothai, S. Asseng, J. Biggs, F. Bert, P.  
20  
21 1038 Bertuzzi, V. Bhatia, M. Bindi, I. Broad, D. Cammarano, R. Carretero, A.A. Chattha, U.  
22  
23  
24 1039 Chung, S. Debats, P. Deligios, G. De Sanctis, T. Dhliwayo, B. Dumont, L. Estes, F.  
25  
26 1040 Ewert, R. Ferrise, T. Gaiser, G. Garcia, S. Gbegbelegbe, V. Geethalakshmi, E.  
27  
28  
29 1041 Gerardeaux, R. Goldberg, B. Grant, E. Guevara, J. Hickman, H. Hoffmann, H. Huang, J.  
30  
31 1042 Hussain, F.B. Justino, A.S. Karunaratne, A.-K. Koehler, P.K. Kouakou, S.N. Kumar, A.  
32  
33  
34 1043 Lakshmanan, M. Lieffering, X. Lin, Q. Luo, G. Magrin, M. Mancini, F.R. Marin, A.D.  
35  
36 1044 Marta, Y. Masutomi, T. Mavromatis, G. McLean, S. Meira, M. Mohanty, M. Moriondo,  
37  
38 1045 W. Nasim, N. Negm, F. Orlando, S. Orlandini, I. Ozturk, H.M.S. Pinto, G. Podesta, Z.  
39  
40  
41 1046 Qi, J. Ramarohetra, H.H. Rahman, H. Raynal, G. Rodriguez, R. Rötter, V. Sharda, L.  
42  
43 1047 Shuo, W. Smith, V. Snow, A. Soltani, K. Srinivas, B. Sultan, D.K. Swain, F. Tao, K.  
44  
45  
46 1048 Tesfaye, M.I. Travasso, G. Trombi, A. Topaj, E. Vanuytrecht, F.E. Viscarra, S.A. Wajid,  
47  
48 1049 E. Wang, H. Wang, J. Wang, E. Wijekoon, L. Byun-Woo, Y. Xiaoguang, B.H. Young,  
49  
50  
51 1050 J.I. Yun, Z. Zhao, and L. Zubair, 2015: The AgMIP Coordinated Climate-Crop Modeling  
52  
53 1051 Project (C3MP): Methods and protocols. In *Handbook of Climate Change and*  
54  
55 1052 *Agroecosystems: The Agricultural Model Intercomparison and Improvement Project*  
56  
57 (AgMIP). C. Rosenzweig, and D. Hillel, Eds., ICP Series on Climate Change Impacts,  
58 1053  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 1054           Adaptation, and Mitigation Vol. 3. Imperial College Press, 191-220,  
5  
6  
7 1055           doi:10.1142/9781783265640\_0008.  
8  
9 1056    Nendel, C., M. Berg, K. Kersebaum, W. Mirschel, X. Specka, M. Wegehenkel, K. Wenkel, and  
10  
11  
12 1057           R. Wieland, 2011: The MONICA model: Testing predictability for crop growth, soil  
13  
14 1058           moisture and nitrogen dynamics. Ecological Modelling 222: 1614-1625. doi: DOI  
15  
16 1059           10.1016/j.ecolmodel.2011.02.018.  
17  
18  
19 1060    O'Leary, G., and D. Connor 1996a: A simulation model of the wheat crop in response to water  
20  
21 1061           and nitrogen supply .1. Model construction. Agricultural Systems 52: 1-29. doi:  
22  
23  
24 1062           10.1016/0308-521X(96)00003-0.  
25  
26 1063    O'Leary, G., and D. Connor, 1996b: A simulation model of the wheat crop in response to water  
27  
28  
29 1064           and nitrogen supply .2. Model validation. Agricultural Systems 52: 31-55. doi:  
30  
31 1065           10.1016/0308-521X(96)00002-9.  
32  
33  
34 1066    O'Leary, G., D. Connor, and D. White, 1985: A simulation-model of the development, growth  
35  
36 1067           and yield of the wheat crop. Agricultural Systems 17: 1-26. doi: 10.1016/0308-  
37  
38 1068           521X(85)90019-8.  
39  
40  
41 1069    Olesen, J., B. Petersen, J. Berntsen, S. Hansen, P. Jamieson, and A. Thomsen, 2002: Comparison  
42  
43 1070           of methods for simulating effects of nitrogen on green area index and dry matter growth  
44  
45  
46 1071           in winter wheat. Field Crops Research 74: 131-149. doi: 10.1016/S0378-4290(01)00204-  
47  
48 1072           0.  
49  
50  
51 1073    Palosuo, T., and co-authors, 2011: Simulation of winter wheat yield and its variability in  
52  
53 1074           different climates of Europe: A comparison of eight crop growth models. Europ. J.  
54  
55 1075           Agronomy 35, 103– 114. doi:10.1016/j.eja.2011.05.001  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65



1  
2  
3  
4 1076 Pirttioja, N., T.R. Carter, S. Fronzek, M. Bindi, H. Hoffmann, T. Palosuo, M. Ruiz-Ramos, F.  
5  
6  
7 1077 Tao, M. Trnka, M. Acutis, S. Asseng, P. Baranowski, B. Basso, P. Bodin, S. Buis, D.  
8  
9 1078 Cammarano, P. Deligios, M.F. Destain, B. Dumont, F. Ewert, R. Ferrise, L. François, T.  
10  
11  
12 1079 Gaiser, P. Hlavinka, I. Jacquemin, K.C. Kersebaum, C. Kollas, J. Krzyszczak, I.J. Lorite,  
13  
14 1080 J. Minet, M.I. Minguéz, M. Montesino, M. Moriondo, C. Müller, C. Nendel, I. Öztürk, A.  
15  
16 1081 Perego, A. Rodríguez A.C., Ruane, F. Ruget, M. Sanna, M.A. Semenov, C. Slawinski, P.  
17  
18  
19 1082 Stratonovitch, I. Supit, K. Waha, E. Wang, L. Wu, Z. Zhao, and R.P. Rötter, 2015: A  
20  
21 1083 crop model ensemble analysis of temperature and precipitation effects on wheat yield  
22  
23  
24 1084 across a European transect using impact response surfaces. *Climate Research* 65: 87–105.  
25  
26 1085 doi:10.3354/cr01322  
27  
28  
29 1086 Porter, J.R., and M.A. Semenov. 1995: Non-linearity in climate change impact assessments.  
30  
31 1087 *Journal of Biogeography*, 22(4/5): 597-600.  
32  
33  
34 1088 Porter, J.R., and M.A. Semenov. 2005: Crop responses to climatic variation. *Philosophical*  
35  
36 1089 *Transactions of the Royal Society of London B Biological Sciences* 360: 2021-2035.  
37  
38  
39 1090 Porter, J.R., L. Xie, and Coauthors, 2014: Food Security and Food Production Systems. Chapter  
40  
41 1091 7 of the Working Group 2 contribution to the Intergovernmental Panel on Climate  
42  
43 1092 Change Fifth Assessment Report. 82pp.  
44  
45  
46 1093 Priesack, E., S. Gayler, and H. Hartmann, 2006: The impact of crop growth sub-model choice on  
47  
48 1094 simulated water and nitrogen balances. *Nutrient Cycling in Agroecosystems* 75: 1-13.  
49  
50  
51 1095 doi: DOI 10.1007/s107500-006-9006-1.  
52  
53 1096 Ray, D.K., J.S. Gerber, G.K. MacDonald, and P.C. West, 2015: Climate variation explains a  
54  
55 1097 third of global crop yield variability, *Nat Commun*, 6, doi: 10.1038/ncomms6989.  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1098 Rienecker, M.M., M.J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M.G. Bosilovich,  
1099 S.D. Schubert, L. Takacs, G.K. Kim, S. Bloom, J.Y. Chen, D. Collins, A. Conaty, A. da  
1100 Silva, W. Gu, J. Joiner, R.D. Koster, R. Lucchesi, A. Molod, T. Owens, S. Pawson, P.  
1101 Pegion, C.R. Redder, R. Reichle, F.R. Robertson, A.G. Ruddick, M. Sienkiewicz, and J.  
1102 Woollen, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and  
1103 Applications. *J. Climate*, 24 (14), 3624–3648, doi: 10.1175/JCLI-D-11-00015.1.

1104 Ritchie, J.T., D.C. Godwin, and S. Otter-Nacke, 1985: CERES-wheat: A user-oriented wheat  
1105 yield model. Preliminary documentation.

1106 Ritchie, S., H. Nguyen, and A. Holaday, 1987: Genetic diversity in photosynthesis and water-use  
1107 efficiency of wheat and wheat relatives. *Journal of Cellular Biochemistry*: 43-43.

1108 Rosenzweig, C., J.W. Jones, J.L. Hatfield, J.M. Antle, A.C. Ruane, and C.Z. Mutter, 2015: The  
1109 Agricultural Model Intercomparison and Improvement Project: Phase I activities by a  
1110 global community of science. In *Handbook of Climate Change and Agroecosystems: The  
1111 Agricultural Model Intercomparison and Improvement Project (AgMIP)*. C. Rosenzweig,  
1112 and D. Hillel, Eds., ICP Series on Climate Change Impacts, Adaptation, and Mitigation  
1113 Vol. 3. Imperial College Press, 3-24, doi:10.1142/9781783265640\_0001.

1114 Rosenzweig, C., J.W. Jones, J.L. Hatfield, A.C. Ruane, K.J. Boote, P. Thorburn, J.M. Antle,  
1115 G.C. Nelson, C. Porter, S. Janssen, S. Asseng, B. Basso, F. Ewert, D. Wallach, G.  
1116 Baigorria, J.M Winter. 2013. The Agricultural Model Intercomparison and Improvement  
1117 Project (AgMIP): Protocols and pilot studies. *Agric. Forest Meteorol.* 170, 166-182.  
1118 <http://dx.doi.org/10.1016/j.agrformet.2012.09.011>

1  
2  
3  
4 1119 Rosenzweig, C., J. Elliott, D. Deryng, A.C. Ruane, C. Müller, et al., 2014: Assessing agricultural  
5  
6  
7 1120 risks of climate change in the 21st century in a global gridded crop model  
8  
9 1121 intercomparison. *Proc. Natl. Acad. Sci.*, 111, 3268-3273, doi:10.1073/pnas.1222463110.  
10  
11  
12 1122 Rötter, R.P., T.R. Carter, J.E. Olesen, and J.R. Porter, 2011: Crop-climate models need an  
13  
14 1123 overhaul. *Nature Climate Change*, 1(4): 175-177.  
15  
16 1124 Rötter, R.P., T. Palosuo, K.-C. Kersebaum, C. Angulo, M. Bindi, F. Ewert, R. Ferrise, P.  
17  
18  
19 1125 Hlavinka, M. Moriondo, J.E. Olesen, J. Takáč, and M. Trnka: 2012. Simulation of spring  
20  
21 1126 barley yield in different climatic zones of Northern and Central Europe: A comparison of  
22  
23  
24 1127 nine crop growth models. *Field Crops Research*, 133, 23-36.  
25  
26 1128 Ruane, A.C., S. McDermid, C. Rosenzweig, G.A. Baigorria, J.W. Jones, C.C. Romero, and L.D.  
27  
28  
29 1129 Cecil, 2014: Carbon–Temperature–Water change analysis for peanut production under  
30  
31 1130 climate change: a prototype for the AgMIP Coordinated Climate-Crop Modeling Project  
32  
33 1131 (C3MP). *Global Change Biology*, **20**, 394–407, doi: 10.1111/gcb.12412  
34  
35  
36 1132 Semenov, M.A., and P.R. Shewry, 2011: Modelling predicts that heat stress, not drought, will  
37  
38 1133 increase vulnerability of wheat in Europe. *Nature Scientific Reports* 1:66, doi:  
39  
40  
41 1134 10.1038/srep00066  
42  
43 1135 Senthilkumar, K., P. Bindraban, W. de Boer, N. de Ridder, T. Thiyagarajan, and K. Giller, 2009:  
44  
45  
46 1136 Characterising rice-based farming systems to identify opportunities for adopting water  
47  
48 1137 efficient cultivation methods in Tamil Nadu, India. *Agricultural Water Management* 96:  
49  
50  
51 1138 1851-1860. doi: 10.1016/j.agwat.2009.08.007.  
52  
53 1139 Shibu, M., P. Leffelaar, H. van Keulen, and P. Aggarwal, 2010: LINTUL3, a simulation model  
54  
55 1140 for nitrogen-limited situations: Application to rice. *European Journal of Agronomy* 32:  
56  
57  
58 1141 255-271. doi: DOI 10.1016/j.eja.2010.01.003.  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 1142 Singels, A., M. Jones, F. Marin, A.C. Ruane, and P. Thorburn, 2013: Predicting climate change  
5  
6 1143 impacts on sugarcane production at sites in Australia, Brazil and South Africa using the  
7  
8  
9 1144 Canegro model. Sugar Tech, doi: 10.1007/s12355-013-0274-1.  
10  
11 1145 Singh, P.J., C. Justice, J. Soares, O. Leo, P. Kosuth, I. Jarvis, D. Williams, W. Bingfang, J.  
12  
13  
14 1146 Latham, and I. Becker-Reshef, 2012: GEO-GLAM: A GEOSS-G20 initiative on Global  
15  
16 1147 Agricultural Monitoring. In 39th COSPAR Scientific Assembly. Held 14-22 July 2012, in  
17  
18  
19 1148 Mysore, India. Abstract D2.2-9-12, Vol. 39, p. 1451. Retrieved from  
20  
21 1149 <http://adsabs.harvard.edu/>  
22  
23  
24 1150 Spitters, C.J.T., and A.H.C.M. Schapendonk, 1990: Evaluation of breeding strategies for drought  
25  
26 1151 tolerance in potato by means of crop growth simulation. Plant and Soil 123: 193-203.  
27  
28  
29 1152 Stackhouse, Jr., P.W., S.K. Gupta, S.J. Cox, J.C. Mikovitz, T. Zhang, and L.M. Hinkelman,  
30  
31 1153 2011: The NASA/GEWEX Surface Radiation Budget Release 3.0: 24.5-Year Dataset.  
32  
33 1154 GEWEX News, 21, No. 1, February, 10-12.  
34  
35  
36 1155 Steduto P., T. Hsiao, D. Raes, and E. Fereres, 2009: AquaCrop-The FAO Crop Model to  
37  
38 1156 Simulate Yield Response to Water: I. Concepts and Underlying Principles. Agronomy  
39  
40  
41 1157 Journal 101: 426-437. doi: DOI 10.2134/agronj2008.0139s.  
42  
43 1158 Stenger, R., E. Priesack, G. Barkle, and C. Sperr, 1999: Expert-N A tool for simulating nitrogen  
44  
45  
46 1159 and carbon dynamics in the soil-plant-atmosphere system. Land Treatment collective  
47  
48 1160 proceedings Technical Session, New Zealand.  
49  
50  
51 1161 Stockle, C., M. Donatelli, and R> Nelson, 2003: CropSyst, a cropping systems simulation model.  
52  
53 1162 European Journal of Agronomy 18: 289-307. doi: PII S1161-0301(02)00109-0.  
54  
55 1163 Supit, I., A.A. Hooijer, and C.A. van Diepen, 1994: System Description of the WOFOST 6.0  
56  
57  
58 1164 Crop Simulation Model Implemented in CGMS, Volume 1: Theory and Algorithms.  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

EUR 15956 EN, Joint Research Center, Commission of the European Communities, Luxembourg.

Tao F., and Z. Zhang, 2010: Adaptation of maize production to climate change in North China Plain: Quantify the relative contributions of adaptation options. *European Journal of Agronomy* 33: 103-116. doi: DOI 10.1016/j.eja.2010.04.002.

Tao F, and Z. Zhang, 2011: Climate change, wheat productivity and water use in the North China Plain: A new super-ensemble-based probabilistic projection. *Agricultural and Forest Meteorology* ISSN 0168-1923,.

Tao, F., M. Yokozawa, and Z. Zhang, 2009a: Modelling the impacts of weather and climate variability on crop productivity over a large area: A new process-based model development, optimization, and uncertainties analysis. *Agricultural and Forest Meteorology* 149: 831-850. doi: 10.1016/j.agrformet.2008.11.004.

Tao F, Z. Zhang, J. Liu, and M. Yokozawa, 2009b: Modelling the impacts of weather and climate variability on crop productivity over a large area: A new super-ensemble-based probabilistic projection. *Agricultural and Forest Meteorology* 149: 1266-1278. doi: 10.1016/j.agrformet.2009.02.015.

van Diepen, C.A., J. Wolf, and H. van Keulen, 1989: WOFOST: a simulation model of crop production. *Soil Use and Management*, 5: 16-24

Vitart, F., A.W. Robertson, and D.L.T. Anderson, 2012: Subseasonal to seasonal prediction project: Bridging the gap between weather and climate. *WMO Bulletin*, 61(2), 23-28

Waha, K., L.G.J. van Bussel, , C. Müller, and A. Bondeau, 2012: Climate-driven simulation of global crop sowing dates. *Global Ecology and Biogeography*, 21: 247–259. doi:10.1111/j.1466-8238.2011.00678.x

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1188 Wallach, D., L.O. Mearns, M. Rivington, J.M. Antle, and A.C Ruane, 2015: Uncertainty in  
1189 agricultural impact assessment. In Handbook of Climate Change and Agroecosystems:  
1190 The Agricultural Model Intercomparison and Improvement Project (AgMIP). C.  
1191 Rosenzweig, and D. Hillel, Eds., ICP Series on Climate Change Impacts, Adaptation, and  
1192 Mitigation Vol. 3. Imperial College Press, 223-260, doi:10.1142/9781783265640\_0009.

1193 Wang, E., and T. Engel, 2000: SPASS: a generic process-oriented crop model with versatile  
1194 windows interfaces. Environmental Modelling & Software 15: 179-188.

1195 White, J.W., G. Hoogenboom, B.A. Kimball, and G.W. Wall, 2011a: Methodologies for  
1196 simulating impacts of climate change on crop production. Field Crops Research 124:357-  
1197 368.

1198 White, J.W., G. Hoogenboom, P.W. Wilkens, P.W Stackhouse, J.M. Hoel, 2011b: Evaluation of  
1199 Satellite-Based, Modeled-Derived Daily Solar Radiation Data for the Continental United  
1200 States. Agron. J., 103(4): 1242–1251. doi: 10.2134/agronj2011.0038.

1201 Williams J., C. Jones, J. Kiniry, D. Spinel, 1989: The EPIC crop growth-model. Transactions of  
1202 the ASAE 32: 497-511.

1203 Yin, X., and H.H van Laar, 2005: Crop Systems Dynamics: An Ecophysiological Simulation  
1204 Model for Genotype-by-environment Interactions. Wageningen Academic Pub., 155pp.