Are soybean models ready for climate change food impact assessments?

Kritika Kothari, Rafael Battisti, Kenneth Boote, Sotirios Archontoulis, Adriana Confalone, Julie Constantin, Santiago Cuadra, P Debaeke, Babacar Faye, Brian Grant, et al.

HAL Id: hal-03666419
https://hal.inrae.fr/hal-03666419

Submitted on 5 Feb 2024

Are soybean models ready for climate change food impact assessments?

Kritika Kothari
University of Kentucky

Rafael Battisti
Universidade Federal de Goiás, Brazil

Kenneth J. Boote
University of Florida

Sotirios V. Archontoulis
Iowa State University

Adriana Confalone
Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina
Follow this and additional works at: https://uknowledge.uky.edu/pss_facpub

Part of the Plant Sciences Commons, and the Soil Science Commons
See next page for additional authors
Right click to open a feedback form in a new tab to let us know how this document benefits you.

Are soybean models ready for climate change food impact assessments?

Authors

Kritika Kothari, Rafael Battisti, Kenneth J. Boote, Sotirios V. Archontoulis, Adriana Confalone, Julie Constantin, Santiago V. Cuadra, Philippe Debaeke, Babacar Faye, Brian Grant, Gerrit Hoogenboom, Qi Jing, Michael van der Laan, Fernando Antônio Macena da Silva, Fabio R. Marin, Alireza Nehbandani, Claas Nendel, Larry C. Purcell, Budong Qian, Alex C. Ruane, Céline Schoving, Evandro H. F. M. Silva, Ward Smith, Afshin Soltani, Amit Srivastava, Nilson A. Vieira Jr., Stacey Slone, and Montserrat Salmerón

Are soybean models ready for climate change food impact assessments?

Kritika Kothari [a], Rafael Battisti [b], Kenneth J. Boote [c], Sotirios V. Archontoulis [d], Adriana Confalone [e], Julie Constantin [f], Santiago V. Cuadra [g], Philippe Debaeke [f], Babacar Faye [h], Brian Grant [i], Gerrit Hoogenboom [c,j], Qi Jing [i], Michael van der Laan [k], Fernando Antônio Macena da Silva [g], Fabio R. Marin [l], Alireza Nehbandani [m], Claas Nendel [n,o], Larry C. Purcell [p], Budong Qian [i], Alex C. Ruane [q], Céline Schoving [f,t], Evandro H.F.M. Silva [l], Ward Smith [i], Afshin Soltani [m], Amit Srivastava [r], Nilson A. Vieira Jr. [l], Stacey Slone [s], Montserrat Salmerón [a,*]

[a] *University of Kentucky, Department of Plant and Soil Sciences, Lexington, KY, United States*
[b] *Universidade Federal de Goiás, Escola de Agronomia, Goiânia, Brazil*
[c] *University of Florida, Agricultural and Biological Engineering Department, Gainesville, FL, United States*
[d] *Iowa State University, Department of Agronomy, Ames, IA, United States*
[e] *Universidad Nacional del Centro de la Provincia de Buenos Aires, Facultad de Agronomía, NACT-NAACCE, Azul, Buenos Aires, Argentina*
[f] *Université de Toulouse, INRAE, UMR AGIR, Castanet-Tolosan, France*
[g] *Empresa Brasileira de Pesquisa Agropecuária, Brasília, Brazil*
[h] *Institut de recherche pour le développement (IRD) ESPACE-DEV, Montpellier Cedex, France*
[i] *Agriculture and Agri-Food Canada, Ottawa Research and Development Centre, Ottawa, Canada*
[j] *University of Florida, Food Systems Institute, Gainesville, FL, United States*
[k] *University of Pretoria, Department of Plant and Soil Sciences, Pretoria, South Africa*
[l] *University of São Paulo, College of Agriculture "Luiz de Queiroz" – Esalq, Biosystems Engineering Department, São Paulo, Brazil*
[m] *Gorgan University of Agricultural Sciences and Natural Resources, Agronomy Group, Gorgan, Iran*
[n] *Institute of Biochemistry and Biology, University of Potsdam, Potsdam, Germany*
[o] *Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany*
[p] *University of Arkansas, Department of Crop, Soil, and Environmental Sciences, Fayetteville, Arkansas, USA*
[q] *NASA Goddard Institute for Space Studies, New York, NY, United States*
[r] *University of Bonn, Institute of Crop Science and Resource Conservation, Bonn, Germany*
[s] *Dr. Bing Zhang Department of Statistics, University of Kentucky, Lexington, KY, USA*
[t] *Terres Inovia, Station Inter-Instituts, Baziège, France*

ARTICLE INFO

ABSTRACT

An accurate estimation of crop yield under climate change scenarios is essential to quantify our ability to feed a growing population and develop agronomic adaptations to meet future food demand. A coordinated evaluation of yield simulations from process-based eco-physiological models for climate change impact assessment is still missing for soybean, the most widely grown grain legume and the main source of protein in our food chain. In this first soybean multi-model study, we used ten prominent models capable of simulating soybean yield under varying temperature and atmospheric CO_2 concentration [CO_2] to quantify the uncertainty in soybean yield simulations in response to these factors. Models were first parametrized with high quality measured data from five contrasting environments. We found considerable variability among models in simulated yield responses to increasing temperature and [CO_2]. For example, under a + 3 °C temperature rise in our coolest location in Argentina, some models simulated that yield would reduce as much as 24%, while others simulated yield increases up to 29%. In our warmest location in Brazil, the models simulated a yield reduction ranging from a 38% decrease under + 3 °C temperature rise to no effect on yield. Similarly, when increasing [CO_2] from 360 to 540 ppm, the models simulated a yield increase that ranged from 6% to 31%. Model calibration did not reduce variability across models but had an unexpected effect on modifying yield responses to temperature for some of

---

* Corresponding author.
E-mail address: msalmeron@uky.edu (M. Salmerón).

the models. The high uncertainty in model responses indicates the limited applicability of individual models for climate change food projections. However, the ensemble mean of simulations across models was an effective tool to reduce the high uncertainty in soybean yield simulations associated with individual models and their parametrization. Ensemble mean yield responses to temperature and [CO_2] were similar to those reported from the literature. Our study is the first demonstration of the benefits achieved from using an ensemble of grain legume models for climate change food projections, and highlights that further soybean model development with experiments under elevated [CO_2] and temperature is needed to reduce the uncertainty from the individual models.

1. Introduction

Soybean (*Glycine max* L. (Merr.)) is globally the fifth most widely grown crop and first among legumes. Future soybean yield projections are essential to develop management adaptations and inform policy-makers to ensure sufficient and sustainable productivity of this major crop. Uncertainties in crop yield projections that hinder effective planning and adaptation to a changing climate are attributed to multiple sources. First, there is uncertainty associated with different climate scenarios from global climate circulation models that is based on a range of trajectories of future greenhouse gas emissions (Moss et al., 2010; Müller et al., 2021; Qian et al., 2020). Additionally, there is uncertainty associated with the responses from different process-based eco-physiological models that are typically used to estimate crop yield under a changing climate (Müller et al., 2021; Ruane et al., 2016). The Agricultural Model Inter-comparison and Improvement Project (AgMIP) was established to quantify and reduce uncertainties in projections of agricultural production by facilitating multi-model analyses and improvement of existing crop models Rosenzweig et al., 2013). While there have been prior crop multi-model comparison studies in wheat (Asseng et al., 2013), maize (Bassu et al., 2014), rice (Li et al., 2015), sugarcane (Marin et al., 2015), and potato (Fleisher et al., 2017), this type of coordinated effort, comparing responses of crop models to key climate change factors, is still missing for soybean and other legume crops. In a prior assessment of future yield projections from global gridded crop models, there was a higher variability for soybean than for the other major crops (Jägermeyr, 2021; Rosenzweig et al., 2014). Thus, there is a pressing need to address this scientific gap and for a coordinated pilot multi-model evaluation to support global yield projections for soybean.

Differences in model responses to temperature and [CO_2] are main sources of variability in yield simulations under climate change (Asseng et al., 2013). Consequently, an essential first step for a new coordinated multi-model initiative is to evaluate how soybean models differ in their responses to temperature and [CO_2] under contrasting environments and as affected by various levels of input data used for calibration. This approach has proven useful to identify sources of model variability and opportunities to reduce uncertainty in yield simulations (Asseng et al., 2013; Bassu et al., 2014; Fleisher et al., 2017; Li et al., 2015; Marin et al., 2015; Wang et al., 2017). For instance, in prior crop model inter-comparisons, researchers discovered the value of using multi-model ensembles to reduce the uncertainty in yield simulations for various applications (Asseng et al., 2013; Bassu et al., 2014; Fleisher et al., 2017; Li et al., 2015; Marin et al., 2015; Martre et al., 2015). These teams showed that individual models may perform better for the regions and conditions for which they were developed and thoroughly tested. However, many of these models did not reproduce observed yield for all treatments and regions consistently well. Conversely, the model ensemble results, for these different crops, were close to the measured yield, even with limited available data for calibration.

In this study, we evaluated ten soybean models and their multi-model ensemble for the simulation of yield under baseline climate, and the sensitivity of models to variable temperature and [CO_2] scenarios. Our goal was to quantify uncertainties in yield simulations under baseline conditions and under changing climate that may be associated with the use of different models, or with the amount of data available for calibration. An additional goal was to assess the applicability of multi-model ensembles to reduce uncertainty in yield simulations from soybean models. We selected five sites with contrasting climate that had high quality observed data for model evaluation under baseline conditions. Four of the sites represent major global soybean production areas; the USA, Argentina, and Brazil together account for greater than 80% of global soybean production (Grassini et al., 2021) (Fig. 1). The results presented have direct implications toward improving future soybean yield projections and informing policy decision-makers.

2. Methods

2.1. Model initialization and calibration under baseline conditions

We calibrated ten eco-physiological soybean models (Supplementary Tables 1 and 2) using experimental data from five sites (Fig. 1). There were 19 treatments across the five sites, including irrigated, rainfed, and planting date treatments depending on the experimental site (Supplementary Table 3). Briefly, there were irrigated and rainfed treatments in Azul, Argentina for three years (n = 6 treatments) and Auzeville, France for two years (n = 4 treatments); all irrigated treatments in Fayetteville, Arkansas, US for three years (n = 3 treatments); all rainfed treatments with early and late planting dates in Ames, Iowa, US for two years (n=4 treatments) and Brasilia, Brazil for one year (n = 2 treatments). The long-term mean growing season temperature ranged from 19.8 °C (Argentina) to 25 °C (Arkansas, USA) and seasonal precipitation ranged from 253 mm (France) to 640 mm (Brazil) (Fig. 1). Soil information for parametrization of soil physical and chemical properties was provided to the crop modelers (Supplementary Table 4).

Weather data, consisting of daily total solar radiation, maximum and minimum temperature, total precipitation, wind speed, relative humidity, and dew point temperature were obtained from weather stations on-site or close to the experimental sites for model simulations during the experimental years. For the long-term (1980–2010) simulations, data were obtained from weather stations on-site when available, or from the closest weather station at the same location with complete data. Gaps and outliers in the long-term weather data were filled with the AgMERRA dataset (Ruane et al., 2015), adjusted after comparing the monthly values of the observed dataset (see Supplementary Information, SI1).

Management information including planting date, cultivar maturity group, irrigation dates and amounts, were provided to describe the different treatments (Supplementary Table 3). Nitrogen fertilizer was only applied at the Brazil site at a low rate of 8 kg N ha^{-1}. Volumetric soil water content and soil inorganic nitrogen concentration was provided based on measurements or preliminary simulations initialized on January 1 (France, Arkansas, and Iowa), or July 1 (Brazil and Argentina).

Each model was calibrated in two steps using different levels of data for calibration: 1) *Blind* calibration with only phenology data; and 2) *Full* calibration with in-season stem, pod, leaf, total aboveground biomass, seed weight, and leaf area index (LAI). The calibrations followed a similar approach that was used in the other multi-model comparisons (Bassu et al., 2014; Fleisher et al., 2017), where models were evaluated after a *Blind* and a *Full* model calibration. During *Blind* calibration,

modelers optimized crop phenology coefficients for the prediction of observed dates of beginning flowering (R1) and physiological maturity (R7). In addition, modelers optimized cultivar coefficients for the prediction of beginning seed (R5) when available. Under *Full* calibration, modelers optimized all cultivar growth and phenology coefficients as needed to get simulated yield as well as in-season LAI, total biomass, leaf, stem, and pod weight close to the observed. Soil parameters were also modified as needed during *Full* calibration as modelers tried to match simulated volumetric water content with the observed. A detailed description of model calibration is provided in Supplementary Information, SI2.

2.2. Sensitivity analysis

After calibrating each individual model for baseline conditions during the experimental years, sensitivity exercises were conducted with 30 years (1980–2009) of weather data from each site modified for five temperature levels (−3, 0, +3, +6, and +9 $^{\circ}$C variation from baseline), and five atmospheric [CO_2] levels (360, 450, 540, 630, and 720 ppm). The modified temperature levels were created by decreasing or increasing the daily minimum and maximum temperatures of the baseline time-series. For example, to create the level of baseline + 3 $^{\circ}$C, the daily maximum and minimum temperatures were increased by 3 $^{\circ}$C throughout the 30 years. The relative humidity was kept the same as the baseline even when the temperature was altered, which resulted in a shift in dew point temperature and vapor pressure deficit. For the [CO_2] concentration, a single value of 360 ppm was used for all 30 years for the baseline and was replaced by a higher value to create elevated [CO_2] scenarios. All the treatments in the sensitivity analysis used auto-irrigation to avoid the confounding effect of water stress in studying [CO_2] and temperature responses. The auto-irrigation was triggered to refill the soil profile to field capacity when there was 50% or less crop available water in the top 0.3 m of soil. There were no fertilizer

applications, and the crop management during the sensitivity analysis was kept the same as one of the treatments during calibration. All the models simulated biological nitrogen fixation, except AQUACROP. The calibration and sensitivity simulations were carried out by teams for each model listed in Supplementary Table 1. Simulations started on 1 January (France, Arkansas, and Iowa), or 1 July (Brazil and Argentina) of each year, and soybean was sown on 6 June in Arkansas, 11 November in Argentina and Brazil, 2 May in Iowa, and 11 May in France. The management and initial conditions of the sensitivity setup were taken from one of the treatments of calibration (1999 Irrigated in Argentina, 2012 Irrigated in Arkansas, Early Planted Rainfed in Brazil, 2017 Irrigated in France, and 2015 Early Planted Rainfed in Iowa). The management was held constant during the sensitivity analysis.

2.3. Multi-model ensemble

To evaluate multi-model ensembles under baseline conditions, we created a total of 3200 random combinations of multi-model ensembles composed of 1 and up to 10 models in the group. Therefore, any given model was selected randomly with repetition to create the multi-model ensemble using the sample function in the R statistical software version 3.6.1 (Team, 2019). The number of multi-model combinations was set to 3200, since a larger number of combinations was found by Martre et al. (2015) to have no effect on results in a multi-model ensemble from a pool of 27 wheat models. Thus, 3200 model combinations from a pool of 10 soybean models provided robust estimates of the mean and nRMSE in model ensemble predictions in our study. This multi-model ensemble analysis assumed that any combination of models has the same chance of being selected as part of the ensemble. However, in reality model users may force their preferred model to be included in the ensemble, as opposed to creating an ensemble of random models, which may not include the model with which they are more experienced. For this purpose, we created a second type of model ensemble by imposing the
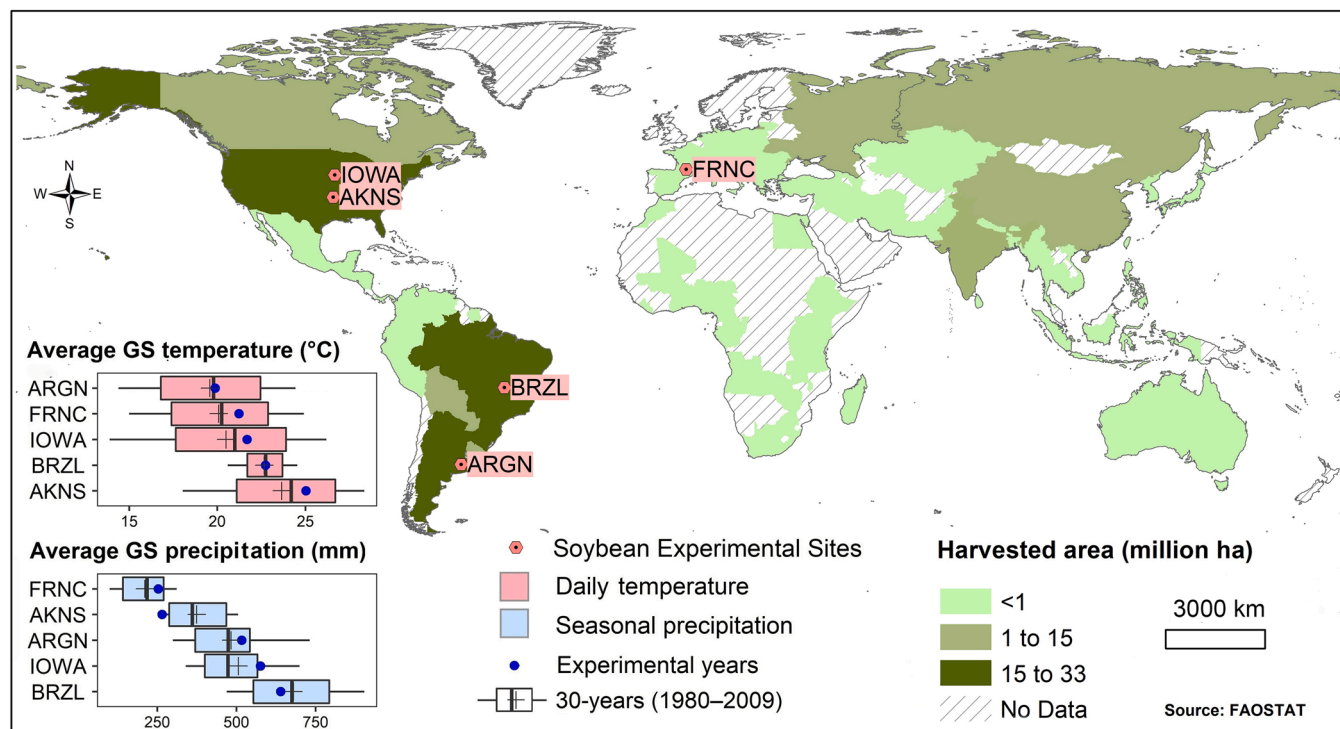
Fig. 1. Locations of experimental sites used for model calibration under baseline conditions and soybean harvested area from 2010 to 2019 (FAOSTAT). Azul, Argentina (ARGN); Brasilia, Brazil (BRZL); Auzeville, France (FRNC); and Ames, IA (IOWA) and Fayetteville, Arkansas (AKNS), USA. Figure inserts show average daily growing season (GS) temperature and total precipitation during the experimental years (blue points), and during 30-years (1980–2009) used for sensitivity analysis (boxplots). Boxplot bounds show 25th and 75th percentiles, whiskers are 10th and 90th percentiles, the median is shown as a vertical bold line, and the mean as a plus symbol.

condition that a particular model must be always present in the model ensemble. With this approach we obtained 10 model ensembles, one for each of the models included in the study, by selecting model combinations that always included a given model from the 3200 random combinations previously generated.

2.4. Data analysis

We tested the efficacy of each individual model as well as a multi-model ensemble in reproducing seed yield and other measured variables using the normalized root mean square error (nRMSE).

$$nRMSE_{m,s} = \sqrt{\frac{\sum_{i=1}^{N_s}\left(Y_{s,i} - \widehat{Y}_{m,s,i}\right)^2}{N_s}} \times \frac{100}{\overline{Y}_s} \qquad (1)$$

where $nRMSE_{m,s}$ is the nRMSE for the model or multi-model ensemble combination m at site s; $Y_{s,i}$ is the ith measured value at site s, and $\widehat{Y}_{m,s,i}$ is the corresponding simulated value by model m; $\overline{Y}_s$ is the average of measured values at site s; and N_s is the number of measured values at site s. The nRMSE was calculated for each model and each site.

The yield change (%) from the baseline under either a $+ 3\ ^\circ C$ or a 540 ppm $[CO_2]$ was analyzed with analysis of variance (ANOVA) using the PROC MIXED procedure in the SAS 9.4 software. Crop model, calibration (*Blind* or *Full*), site, and their interactions were considered fixed factors in the ANOVA model. The year nested within site and the interaction of crop model with year nested within site were considered as random effects. We calculated the yield change from the baseline before analysis by dividing simulated yield simulations within each year and level by yield simulations under the baseline ($+0\ ^\circ C$ temperature variation at 360 ppm $[CO_2]$) and expressing the yield change as a percentage. We quantified the effect of calibration on the yield change (%) using a least significance means of the crop model, location, and calibration interaction with a slice by crop model and location. We transformed the p-value using the LogWorth statistic (Greenland, 2019; Sall, 2002) (LW), defined as $-\log10$ (p-value), to compare the effect of calibration in different models given the high incidence of $p < 0.001$ in our results due to crop models being deterministic in nature and void of random errors. To overcome the sensitivity of p-value in the large dataset and its favoring of factors with multiple levels, the LW statistic was found more suitable for testing calibration effect in our study (Greenland, 2019; Sall, 2002).

In addition to analyzing the percent yield and biomass change under a $3\ ^\circ C$ temperature increase, we investigated the relationship between these variables and mean growing season temperature utilizing data from model simulations at temperature levels ranging from 0 to $+ 9\ ^\circ C$. The $- 3\ ^\circ C$ temperature level was excluded from this analysis due to crop development rate being slowed down causing crop failure or causing steep changes in the shape of the response for some of the years. We first normalized grain yield and biomass by year, location, model, and calibration (See more details on the normalization approach in the Supplementary Information, SI3). Once yield data were normalized, we quantified the effect of model and calibration on the linear or quadratic relative yield or biomass response to mean growing season temperature with an analysis of covariance (ANCOVA) using the PROC MIXED procedure in SAS. Model, calibration, and site were considered as fixed factors, and mean growing season temperature and its square were covariates. The quadratic term was dropped from the model when not significant at $P < 0.005$ within a given location, model, and calibration. The shape of the relationship was specified for each location in the ANCOVA model by multiplying the square term of the model by a constant with a value of 0 or 1. The resulting ANCOVA model (Supplementary Table 9) was used to quantify the effect of calibration on the shape of the response within each model and location with a contrast statement and expressed the p-value using the LW statistic, the higher the LW the lesser the probability of rejecting that the calibration effect is

significant. The same approach was used to analyze relative biomass responses to mean growing season temperature by location, model, and calibration.

We compared our model-simulated yield response to an increase in temperature with literature reported values from elevated temperature experiments conducted in outdoor heated experiments (Burkey et al., 2020; Ruiz-Vera et al., 2013), controlled environment (Alsajri et al., 2020; Xu et al., 2016), and temperature gradient chambers (Baek et al., 2020; Tacarindua et al., 2013). In order to normalize the yield-temperature responses from the literature and our simulations, we calculated the yield change per $1\ ^\circ C$ increase in temperature and chose studies that increased temperature higher than $1.5\ ^\circ C$ compared to the control treatment. Similarly, we compared the simulated yield response to elevated $[CO_2]$ from our study with literature-reported values from Free-Air Carbon dioxide Enrichment (FACE) experiments (Bishop et al., 2015; Gray et al., 2016; Hao et al., 2014; Morgan et al., 2005; Ruiz-Vera et al., 2013) and controlled environment experiments (Baker et al., 1989; Wang et al., 2018). For a fair comparison, we compared our simulated yield increase from baseline to elevated $[CO_2]$ of 540 ppm with experiments where $[CO_2]$ was increased from baseline to 540–600 ppm. Likewise, we compared yield response to $[CO_2]$ increased to 720 ppm in our study with experiments where $[CO_2]$ was raised to 600–800 ppm. Further details regarding the experimental setup from the literature review are provided in supplementary section SI4.

3. Results

### 3.1. Baseline conditions and effect of calibration

Measured soybean seed yield (Fig. 1) ranged from 3297 to 6179 kg ha^{-1} depending on the site and treatment (Fig. 2a). After *Full* calibration with detailed in-season observations, the variability and normalized root mean square error (nRMSE) for the simulated yield across the ten models was reduced for each site compared to the *Blind* calibration that was based on phenological observations only (Fig. 2b). Individual models simulated yield across all sites with an nRMSE that ranged from 20% to 51% after *Blind* calibration, and from 9% to 31% after *Full* calibration (Fig. 3a). As expected, the nRMSE for simulation of in-season biomass was also reduced from 22% to 109% under *Blind* calibration, to 16–48% across models after *Full* calibration (Supplementary Fig. 1). Similarly, the nRMSE for simulation of LAI was reduced from 27% to 83% across models under *Blind* calibration, to 25–69% under *Full* calibration (Supplementary Fig. 1). The yield nRMSE after *Blind* calibration was greater than the 13.3% coefficient of variation (CV) typical across replications in soybean variety trials (Storck et al., 2010), while after *Full* calibration four models simulated yield with an nRMSE below 13.3%. The ensemble mean and median ranked 3rd to 4th for the prediction of seed yield and, thus, were not the best predictors but did have a yield nRMSE below 13.3% (Fig. 3a).

3.2. Number of models required in an ensemble under baseline conditions

We explored the minimum number of models required in the ensemble to reduce the uncertainty for the simulation of yield (Fig. 3b). The nRMSE of the ensemble-mean decreased from 38% to 31% going from one to four models after *Blind* calibration, and from 18% to 14% after *Full* calibration (Fig. 3b). Adding more than four models to the ensemble reduced the nRMSE by less than 1% in absolute units. We also investigated a practical implication of using multi-model ensembles that had not been addressed in the past: modelers may prefer to always keep in the ensemble the model they are most experienced with. We investigated if the number of models required in the ensemble was influenced for ensemble cases containing the best- and worst-performing models within each location (Fig. 4). When the best performing models were always in the ensemble for a given site, adding more models increased the nRMSE. When the worst performing models were always in the
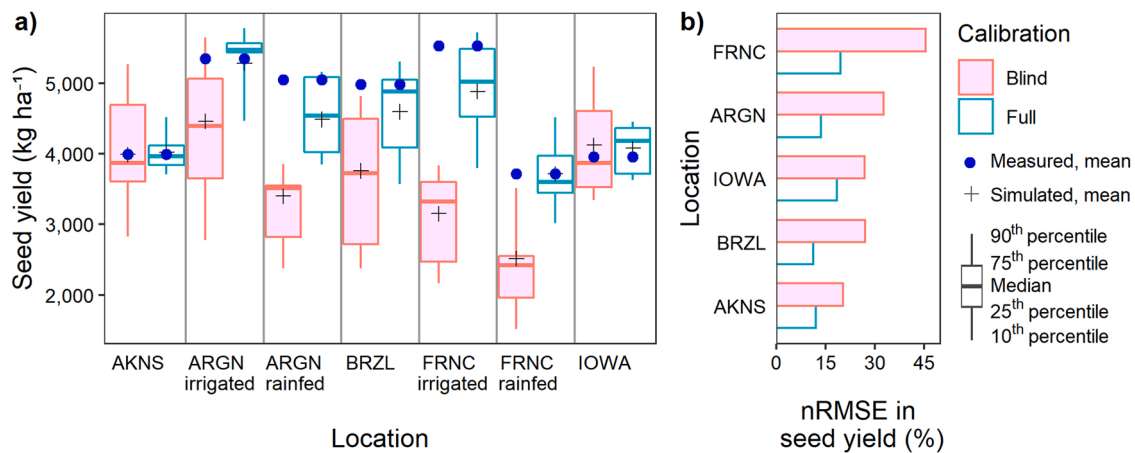
Fig. 2. Mean observed and simulated yield and normalized root mean square error (nRMSE) for simulation of yield at each location after *Blind* and *Full* calibration. (a) Measured (blue symbol) and simulated seed yield across ten models (boxplots) with data averaged across treatments within a location and irrigation system (rainfed or irrigated), and (b) average nRMSE across models for simulation of seed yield by location after *Blind* and *Full* calibration.
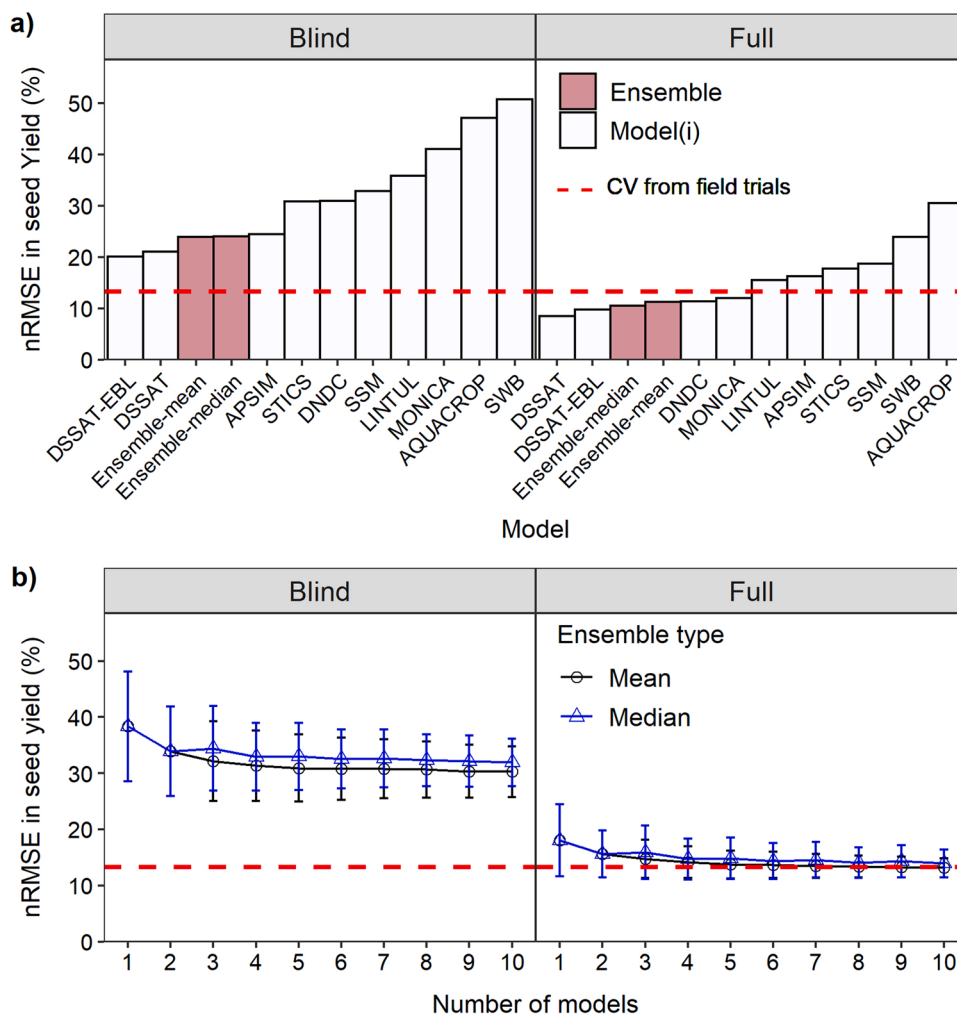
Fig. 3. Normalized root mean square error (nRMSE) for the simulation of seed yield by individual models and for ensembles created with different number of models. (a) Ranking in nRMSE of individual models and ensembles created from the mean and median across all models under Blind and Full calibration (b) nRMSE in ensembles created with an increasing number of models selected randomly with repetition. Error bars represent the standard deviation from 3200 random selections of models that comprise an ensemble. The horizontal dashed line indicates a 13.3% yield co-efficient of variation (CV) obtained from soybean variety tests (Storck et al., 2010).

ensemble, adding more models decreased the nRMSE (Fig. 4). Interestingly, there was no single model that minimized nRMSE at all sites, but different models were best at different sites. In addition, calibration altered the ranking of models within a given site (Fig. 4b).

3.3. Model response to a + 3℃ temperature increase

The models differed greatly in their responses to deviations from the baseline temperature of − 3, 0, + 3, + 6 and 9 °C (Fig. 5a). The ANOVA of + 3 °C temperature change showed that the variability in yield change (%) was mainly attributed to model and location main effects,
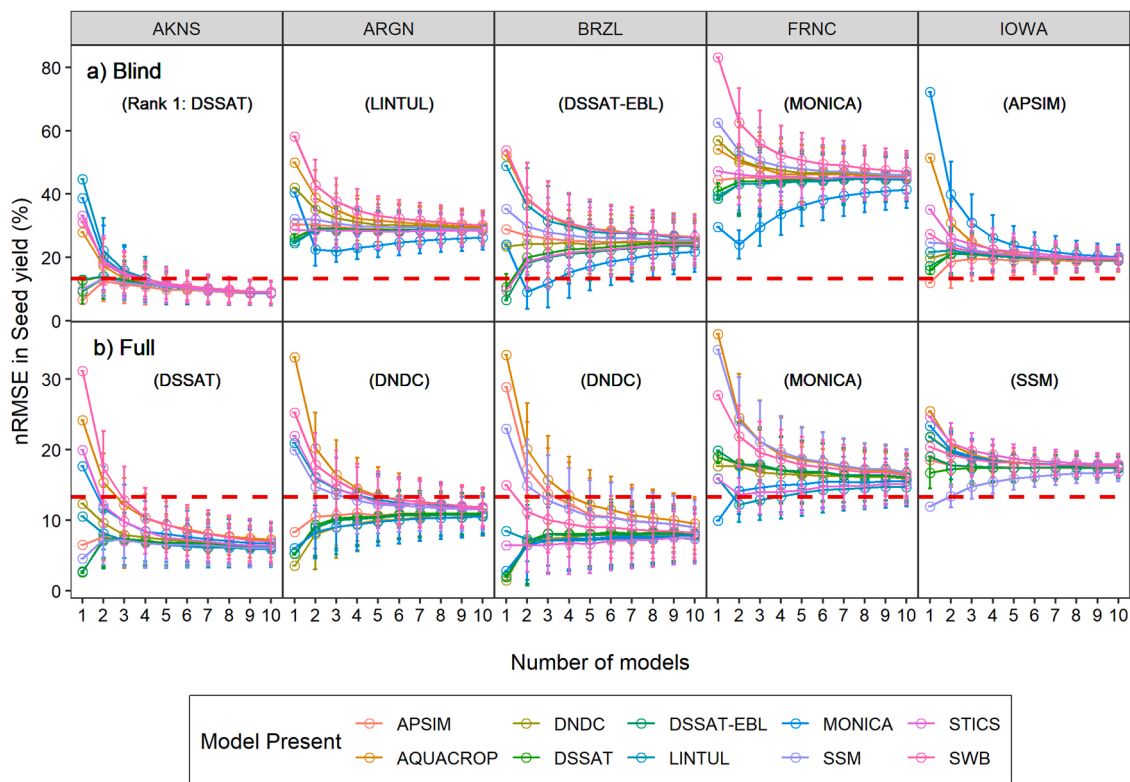
Fig. 4. Normalized Root Mean Square Error (nRMSE) for simulation of observed seed yield imposing the condition that a given model is always present in the ensemble. (a) nRMSE after *Blind* and (b) nRMSE after *Full* calibration at each location by a mean multi-model ensemble composed of an increasing number of models. Rank 1 in parenthesis indicates the model with the lowest nRMSE for the prediction of observed yield within a location. The horizontal dashed line indicates 13.3% CV typical from field trials (Storck et al., 2010).

and was affected by calibration in case of some models and locations (Supplementary Table 5). Most models and locations had a small calibration effect as seen from smaller LW values (Supplementary Table 6). High LW values (LW>= 4 or p = 0.0001) were found only in 13 out of 54 total cases, and the LW value of model ensemble was in the middle of that of the individual models, indicating that it was efficient in filtering large calibration effects. Some of the unaccounted variability in the ANOVA of the relative yield response to temperature was due to a year effect (Supplementary Table 5), given that growing season temperature varies from year to year within each location. The relative yield change in response to a + 3 °C temperature increase was highly variable across models after *Blind* calibration and this variability persisted after *Full* calibration (e.g., 0–42% decline after *Blind* and +3% (SSM) to − 38% (APSIM) after *Full* in Brazil). After *Full* calibration, for a temperature rise of + 3 °C from the baseline, the yield was simulated to decline by 1–19% depending on the model in France, by 12–21% in Iowa, and by 3–32% in Arkansas. For a + 9 °C temperature deviation from baseline, the model variability increased, with a yield decline ranging from 17% to 72% depending on the model in France, 4–70% in Iowa, 14–97% in Brazil, and 29–84% in Arkansas. In addition, for the coldest site, models did not agree on the direction of the yield change in response to the temperature increase. For instance, the yield change in Argentina under a temperature rise of + 3 °C ranged from a 24% yield decline (SWB) to a 29% yield increase (APSIM), and under a temperature raise of + 9 °C, yield changed from − 67% (SWB) to + 9% (APSIM) (Fig. 6a). Across all locations, the SSM model typically simulated the smallest yield decline with an increase in temperature, while the MONICA model predicted the greatest yield decline in response to an increase in temperature.

The multi-model ensemble proved useful for filtering out the extreme temperature responses of the individual models, while also avoiding large calibration effects on these responses (Fig. 5a, Supplementary Table 6). For a 3 °C increase in temperature after *Full* calibration, the

ensemble predicted a yield decrease of 0.2% (Argentina), 5.9% (Iowa), 7.1% (France), 11.6% (Arkansas), and 16.1% (Brazil) (Fig. 5a). Under the most extreme temperature increase (+9 °C), the ensemble simulated yield decreases of 29.1% (Argentina) and 48.8% (Brazil).

3.4. Analysis of model responses to mean growing season temperature

Models differed in the magnitude and shape of the relationship between simulated yield and mean growing season temperature, which was both quadratic and linear for the same location depending on the model and calibration (see example in Fig. 7 and results from all sites in Supplementary Figs. 2a to 2d). Cooler sites such as Argentina were more likely to show a quadratic or non-linear response (Fig. 7). In contrast, for the warm sites in Arkansas and Brazil, yields responses to temperature showed a declining linear relationship in more instances (Supplementary Figs. 2c and 2d). The shape of biomass-temperature response curve was similar to that of yield-temperature response (Supplementary Figs. 3a to 3e), although the reduction in biomass with an increase in temperature was less in magnitude. For instance, the effect of a 6 °C temperature increase on biomass (12–24% reduction) was less than that simulated for yield (12–34% reduction depending on the location).

The effect of calibration on the shape of the yield and biomass-temperature response curves showed that LW values were lower than four (p = 0.0001 or LW = 4) for most models and location combinations, indicating a significant but small calibration effect (Supplementary Figs. 3a to 3e). However, a greater calibration effect that modified the magnitude or shape of the response was found in six out of the 54 cases, and more often for the colder sites (Argentina and France) compared to the warmer sites. For example, in Argentina site, the APSIM and STICS models simulated an almost linear decline in yield with increasing temperature after *Blind* calibration, but after *Full* calibration a quadratic model with an optimum temperature above the baseline provided a
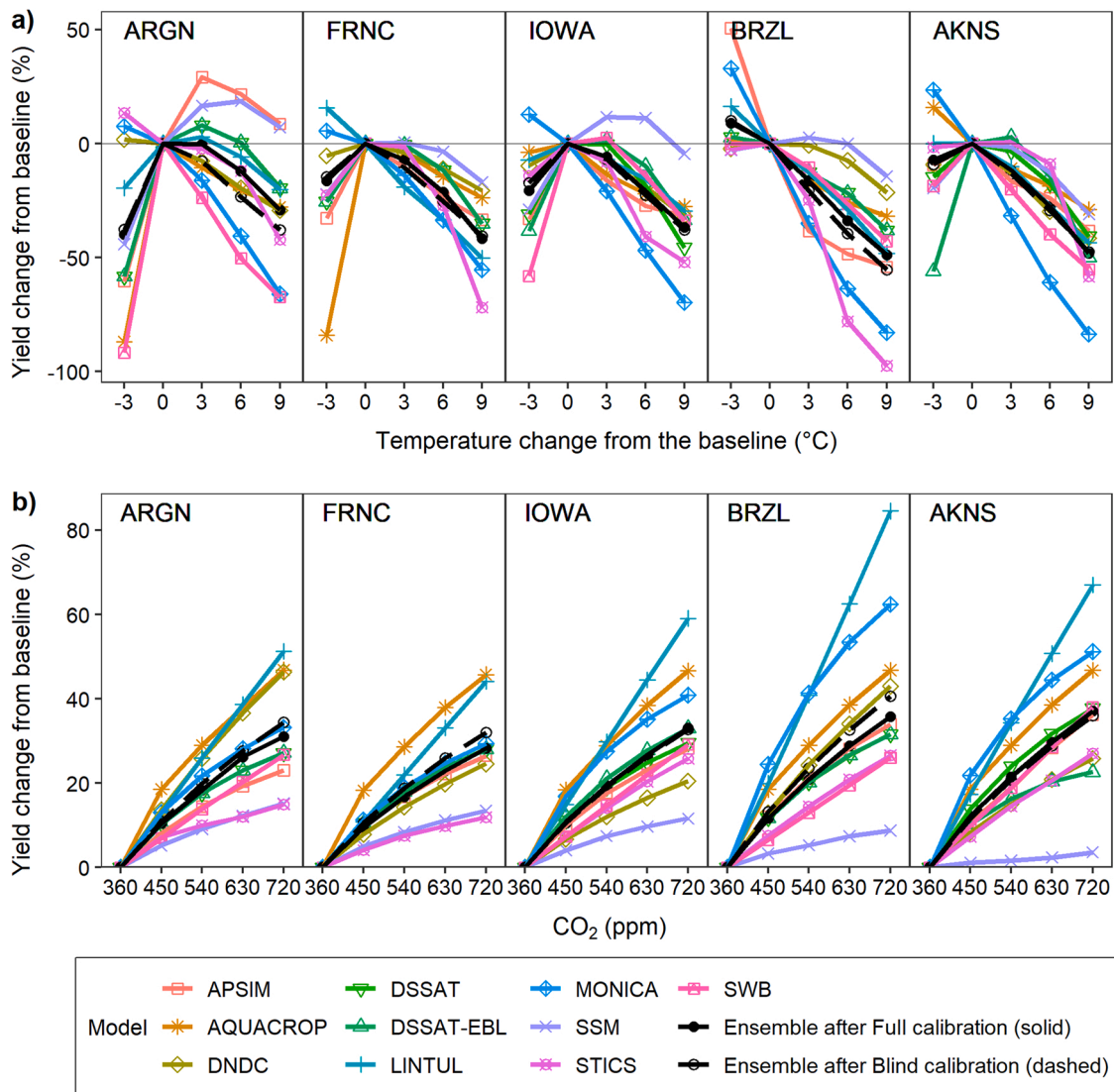
Fig. 5. Simulated yield change (%) in response to a) temperature and b) [CO₂] variation from the baseline at each site. Results are shown for each model after *Full* calibration, and for the model ensemble after *Blind* and *Full* calibration. The locations are sorted from cold (Argentina) to hot (Arkansas), left to right, based on mean growing season temperature. Data averaged across 30 years of simulations under no water limitations (automatic irrigation).

better relative yield response to temperature (Supplementary Figs. 2a to 2d).

The effect of model and calibration on the yield and biomass response curve to temperature may be partially associated with differences in the growing season duration and in the maximum LAI achieved, both of which will influence the amount of photosynthetically active radiation absorbed by the crop canopy. The relationship between the mean growing season temperature and its duration showed that growing season length was generally reduced with increasing temperature in most models (Supplementary Figs. 4a to 4e). However, for a group of models (i.e., APSIM, DSSAT, DSSAT-EBL, and STICS) the relationship between the growing season duration and temperature was best described by a negative quadratic relationship, where growing season length increased at extreme temperatures. The analysis of the relationship between LAI and mean growing season temperature showed that a group of models (i.e., SSM, STICS) simulated an increase in LAI with increasing temperature, another group a decrease in LAI (i.e., APSIM, MONICA, LNTUL), while the remaining models showed a relatively small effect of temperature on maximum LAI (Supplementary Figs. 5a to 5e). Models also differed greatly in their simulation of maximum LAI under baseline conditions after *Blind* calibration e.g., 3.8–12.8 m² m⁻²

maximum LAI across models in Argentina. After *Full* calibration, the difference in simulated LAI decreased in magnitude but still persisted (e. g. 5.1–9.8 m² m⁻² maximum LAI in Argentina).

3.5. Response to [CO₂]

All models simulated a positive yield response to increasing [CO₂], but their responses differed greatly in magnitude (Fig. 5b). The ANOVA of the percent yield increase in response to a [CO₂] increase from 360 to 540 ppm showed that the variations in yield response were mostly attributed to a model effect, whereas calibration and site had a limited effect on this response (Supplementary Tables 5 and 7). The effect of calibration on the yield-[CO₂] response was noticeable for a few model and location combinations; however, the calibration effect was minor on the simulated yield response to [CO₂] by the model ensemble (Supplementary Table 7).

The variation across models in the % yield response to [CO₂] was high after *Blind* calibration, and this variability did not reduce after *Full* calibration (Fig. 8a). For instance, the increase in yield when increasing [CO₂] from 360 to 540 ppm ranged from 12% (STICS) to 29% (LINTUL and AQUACROP) after *Blind* Calibration, and from 6% (SSM) to 31%
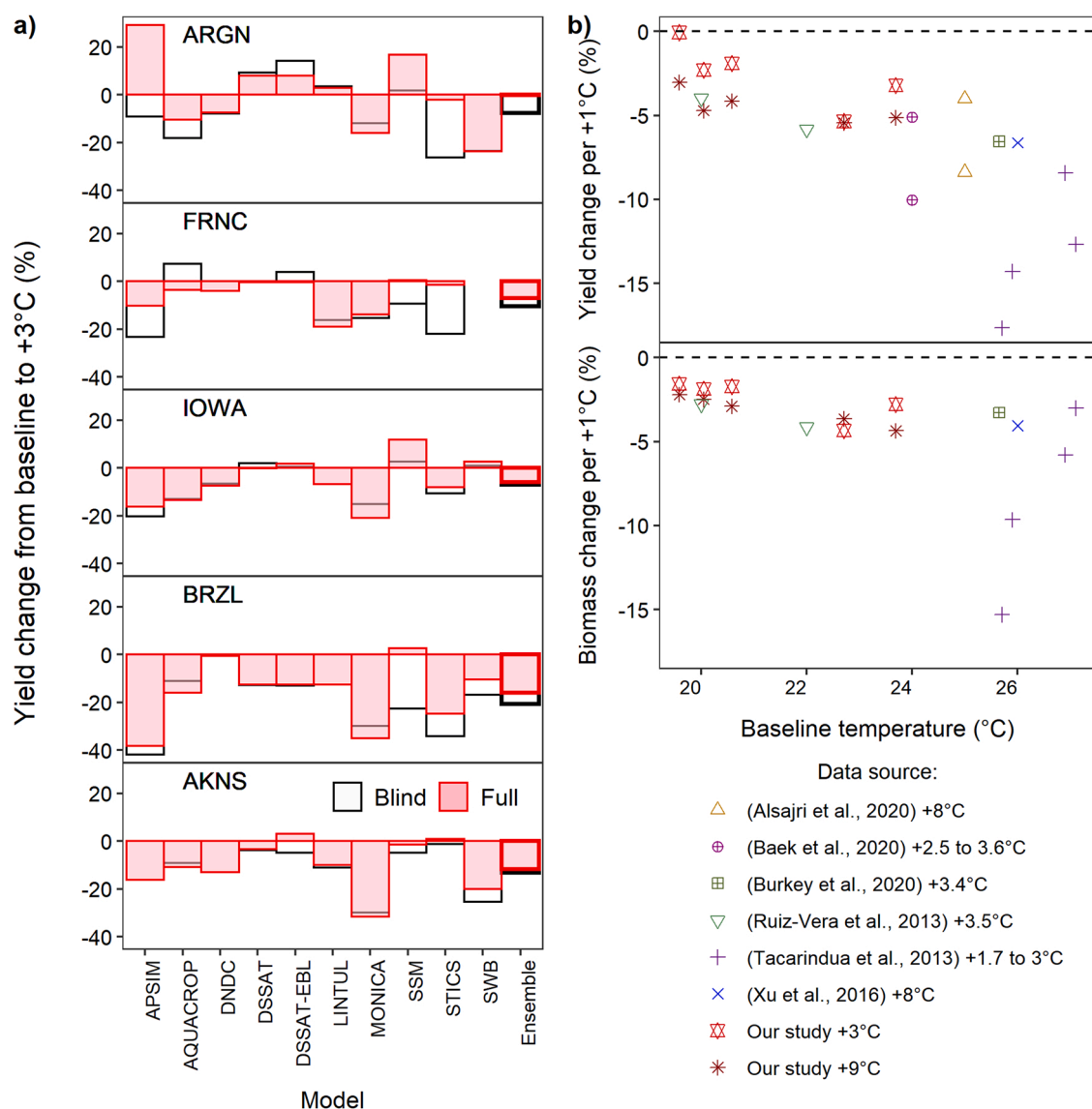
Fig. 6. Yield change under a + 3 °C temperature increase simulated in our study and comparison with a literature review of field experiments. a) Yield change (%) from the baseline by model and calibration for each location based on 30 year simulations. The baseline temperature varied for each location, Argentina (19.8 °C), France (21.2 °C), Iowa (21.7 °C), Brazil (22.7 °C), and Arkansas (25.0 °C). b) Comparison of the rate of yield and biomass change (% change per 1 °C increase in temperature) simulated by the model ensemble and measured in field experiments conducted in out-door heated plots (Burkey et al., 2020; Ruiz-Vera et al., 2013), controlled environment (Alsajri et al., 2020; Xu et al., 2016), and temperature gradient chambers (Baek et al., 2020; Tacarindua et al., 2013) (further details from field experiments are in Supplementary SI4).

(LINTUL) after *Full* calibration (Fig. 8a). Similarly, the yield change when increasing [CO$_2$] from 360 to 720 ranged from 22% (STICS and SSM) to 59% (LINTUL) after *Blind* calibration, and by 11% (SSM) to 61% (LINTUL) after *Full* calibration (Fig. 8a). The LINTUL and AQUACROP models consistently showed the largest yield response to rising [CO$_2$], while the STICS and SSM models had a relatively smaller yield response. When evaluating differences across locations, most models simulated a yield gain under a [CO$_2$] rise from baseline to 540 ppm that was greater for locations with a higher growing season temperature (Fig. 9a). In particular, two models, i.e., LINTUL and MONICA, exhibited a higher yield gain under higher temperature and increased [CO$_2$], while the AQUACROP model showed similar yield responses across locations, and the SSM model showed a negative relationship between yield change and average growing season temperature (Fig. 9 and Supplementary Table 7).

The mean multi-model ensemble was efficient avoiding extreme yield-[CO$_2$] responses from the individual models, with a mean

simulated yield increase ranging from 17% (France) to 22% (Arkansas) for 360–540 ppm after *Full* calibration (Supplementary Table 7). When doubling atmospheric [CO$_2$] from 360 to 720 ppm, the ensemble model simulated a yield increase from 28% to 36% after *Full* calibration. Overall, the simulated yield gain under increasing atmospheric [CO$_2$] had a tendency to be the highest for relatively warm environments such as Brazil and Arkansas, USA (Fig. 9a).

We analyzed the effect of elevated [CO$_2$] on additional model outputs other than yield to identify possible mechanisms for how crop models respond to [CO$_2$] (Supplementary Table 8). For a [CO$_2$] increase from 360 ppm to 540 ppm after *Full* calibration, the simulated increase in yield from the multi-model ensemble (17–23% across sites) was parallel to the simulated gain in total aboveground biomass (17–22%). Thus, the models simulated an increase in photosynthesis and growth under high [CO$_2$] with similar partitioning to reproductive tissues. The maximum LAI also increased with increasing [CO$_2$], indicating that the greater photosynthesis drives LAI in the models (Supplementary Table
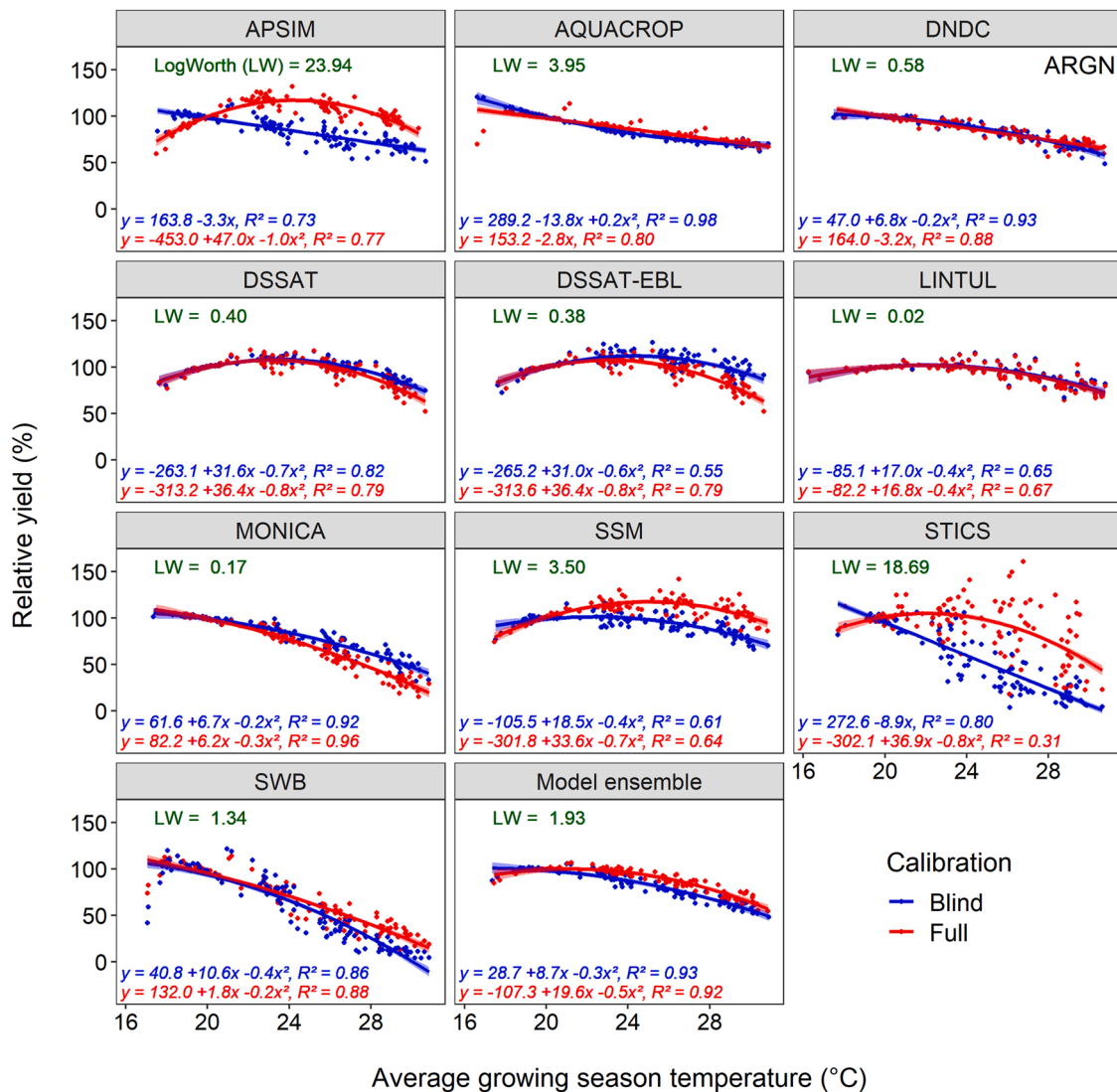
Fig. 7. Relationship between relative yield and mean growing season temperature by model and calibration in Argentina. Data from 30-year simulations and temperature treatments 0 to + 9 °C. The circles show simulated relative yield, while the lines are fitted quadratic or linear curves by model and calibration with equations and R2 shown in figures. The bands are 90% confidence interval of the model fit. LogWorth values are the transformed probability values of the calibration effect within each model, with higher Logworth values indicating a lower probability that the calibration effect is not significant (e.g. LogWorth = 4 is equivalent to p = 0.0001).

8).

4. Discussion

4.1. Baseline conditions

4.1.1. Additional observed field data reduce simulation uncertainty

Model calibration with observed data is a necessary step before conducting field-scale model applications to study management and climate scenarios. However, such task is often impossible at regional-scale simulations due to limited data for calibration (Müller et al., 2017, 2021). Part of soybean yield variability can be explained by the crop growth cycle duration, which allows for good yield estimates from soybean models previously calibrated for a region based on information on cultivar maturity (Salmerón et al., 2017; Setiyono et al., 2010). Similarly, we found that models previously calibrated and evaluated with a larger number of independent soybean experimental data such as DSSAT and APSIM (Archontoulis et al., 2014; Gaydon et al., 2017; Mavromatis et al., 2002; Ruíz-Nogueira et al., 2001), had the lowest nRMSE for yield prediction under *Blind* calibration with phenology data

only (<30%). The APSIM and DSSAT models also showed the lowest nRMSE for simulation of in-season LAI and biomass (<31%) after *Blind* calibration (Supplementary Fig. 1). However, with additional observed data for model calibration, the nRMSE for simulation of in-season biomass and LAI was reduced compared to *Blind* calibration with few exceptions (Supplementary Fig. 1). As a result of calibration with detailed in-season variables and final yield, some of the models with less a priori evaluation had a yield nRMSE that was lower than or similar to the expected experimental error (13.3%) (Storck et al., 2010). Thus, our study provided an opportunity for model improvement of less evaluated soybean models through calibration of cultivar characteristics with detailed crop data.

Another source of uncertainty was associated with the simulation of water balance and water stress. The models showed high variability for simulating the soil and plant water balance as well as sensitivity to drought stress when provided with only phenological data and a local soil characterization (Supplementary Table 4), which contributed to the high nRMSE after *Blind* calibration. After *Blind* calibration, most models tended to under-estimate yield in Argentina and France, especially for rainfed conditions (Fig. 2). Overall, our results suggest that uncertainty
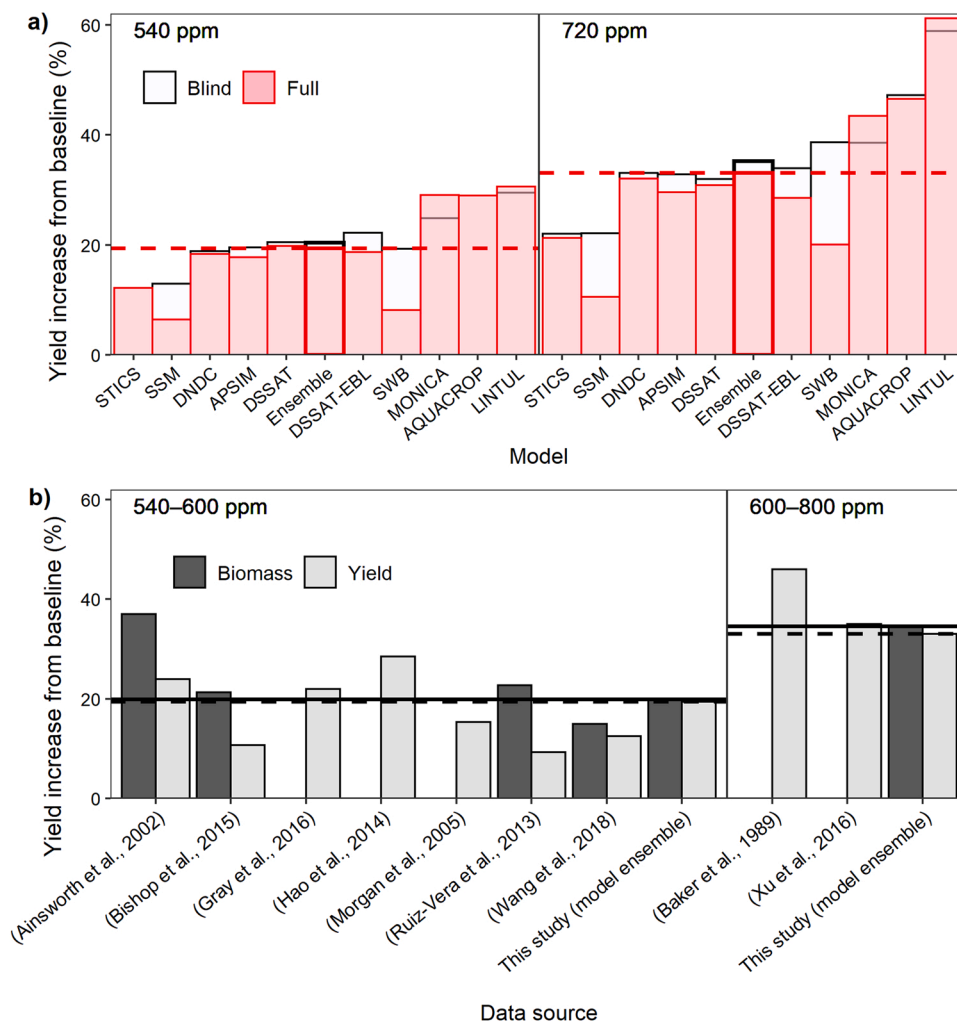
Fig. 8. Yield response under elevated [CO2] levels simulated in our study and comparison with a literature review of field experiments. Yield change (%) from the baseline (360 ppm) under elevated [CO2] levels of a) 540 ppm and b) 720 ppm, obtained from 30-year simulations and averaged across five locations. c) Comparison of yield response simulated by the model ensemble in our study with elevated CO2 experiments in FACE (Bishop et al., 2015; Gray et al., 2016; Hao et al., 2014; Morgan et al., 2005; Ruiz-Vera et al., 2013), controlled environment growth experiments (Baker et al., 1989; Wang et al., 2018), and a meta-analysis of field studies (Ainsworth et al., 2002) (further details from field experiments are in Supplementary SI4). The dashed line across the figure panels shows yield response and the solid line shows biomass response of the model ensemble from our study.

in yield predictions under water stress was associated more to a location effect rather than to the a priori experience of the different soybean models.

4.1.2. Role of multi-model ensembles to manage uncertainty

A model ensemble can reduce uncertainty when limited data for model calibration are available. In prior maize (Bassu et al., 2014), wheat (Wallach et al., 2018), and sugarcane (Marin et al., 2015) multi-model studies, the mean or median across all models were better yield predictors than any individual model. In our study, the ensemble mean or median reduced nRMSE below that of most of the models after *Blind* and *Full* calibration, but they were not the best yield predictors. Prior studies in other crops showed that ensembles can reduce the nRMSE below the yield CV after *Blind* calibration with phenology data (Bassu et al., 2014; Fleisher et al., 2017; Li et al., 2015). In our study, the ensemble was only able to simulate yield with a nRMSE close to a CV of 13.3% obtained in variety trials (Storck et al., 2010) after *Full* calibration (Fig. 3).

We found that adding more than four models to the ensemble under either *Blind* or *Full* calibration did not further reduce the nRMSE (Fig. 3). Similarly, in a wheat multi-model study of 25 models, Wallach et al. (2018) concluded that five models were sufficient to minimize error in yield simulations. Battisti et al. (2017) found that an ensemble of five soybean models had lower bias (262 kg ha^{-1}) than the individual models (535–650 kg ha^{-1}). Model-ensemble analysis considers a random selection of the models that compose an ensemble. However, modelers may prefer to always include the model they are most

experienced with as part of the ensemble. This question had not been addressed in past studies. We found that an ensemble of four or five models was also sufficient to minimize the yield nRMSE when included with the worst-performing model within a site (Fig. 4). Expectedly, an increase in the yield nRMSE was observed when starting with the best-performing model and then adding additional models to the ensemble. No single model minimized nRMSE at all sites, but different models performed best for different sites. Thus, even the best performing models for certain locations, benefited from using a multi-model ensemble when applied to other locations. There are major implications of this finding given that the ranking of models was variable from one site to another, and was affected by calibration. Without a priori information of model performance in a site, it is not wise to assume that a given model is better than a model ensemble. Rather, this finding supports the conclusion of prior multi-model studies (Battisti et al., 2017; Wallach et al., 2018) for use of multi-model ensemble of approximately five models to increase confidence in yield predictions.

Overall, our study found benefits from using a multi-model ensemble that were consistent with prior studies for other crops, but also identified some differences. The relatively high nRMSE of the model ensemble in our study after *Blind* calibration and its modest ranking could be attributed to difficulties in simulating yield under rainfed conditions. This conclusion is consistent with a recent maize study under low input conditions in Africa, where the yield nRMSE was 63% for the *Blind* ensemble and 26% for *Full* ensemble (Falconnier et al., 2020). Another explanation is that soybean models, with the exception of the CROPGRO-Soybean model of DSSAT, have been generally less robustly
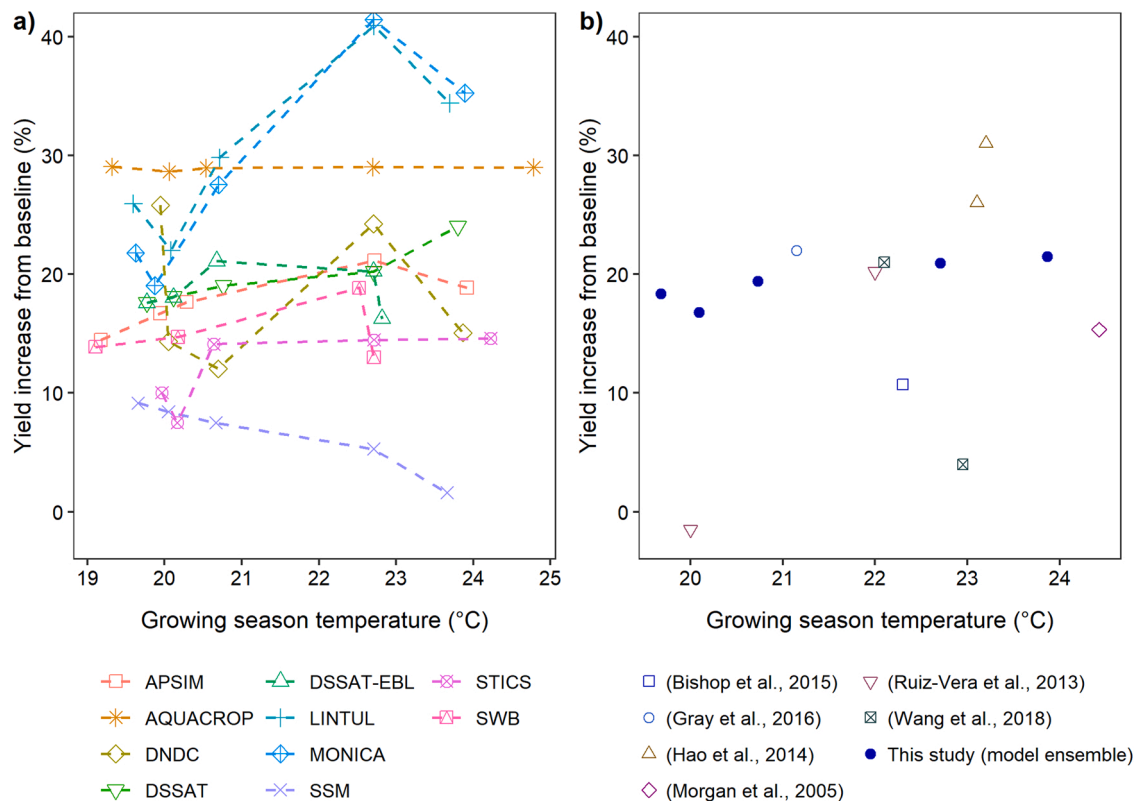
Fig. 9. Yield response under elevated [CO$_2$] as affected by growing season temperature. a) Simulated yield change (%) by individual models at the five different locations in this study; and b) comparison of yield change (%) simulated from the model ensemble in our study with values reported from the literature experimentally. The figure shows the yield change from the baseline (360 ppm) under a [CO$_2$] increase of 540 ppm in our study, and to 540–600 ppm for literature-reported values.

evaluated with experimental data compared to cereal crop models, and in particular evaluated less for processes specific to legumes, including biological N$_2$ fixation (Boote et al., 2009), that is known to be highly sensitive to environmental conditions and drought stress (Purcell et al., 2004).

4.2. Response to temperature

We found a large variability in modeled-yield responses to changes in temperature (Figs. 5 and 6). Under a + 3 °C temperature increase, most models showed a reduction in yield at the warmest locations, but the direction of this effect differed across models at the cooler sites. Different temperature responses at cool versus warm sites were expected because baseline ambient temperatures at each site may be above or below the optimum cardinal temperatures for soybean. For instance, Hatfield et al. (2011) proposed an optimal range of 25–37 °C for vegetative production and 22–24 °C for reproductive seed yield in soybean, with failure of seed yield and harvest index at 39 °C (Battisti et al., 2018; Boote et al., 2005). Thus, an increase in temperature had direct effects on the rate of photosynthesis, growth and addition of reproductive sinks in some models based on their cardinal temperatures for growth and reproductive partitioning (Supplementary Table 2). Daily baseline temperatures were more likely to fall below the range of optimum temperature for soybean photosynthesis and growth (25–30 °C) (Jeffers and Shibles, 1969) for the coldest sites in our study, which may explain the positive yield change that was simulated with + 3 °C temperature increase by some of the models in Argentina.

The relative biomass response to growing season temperature was similar to the relative yield response, but smaller in magnitude. This reduction in harvest index was simulated in many models with a direct temperature effect on the rate of partitioning to reproductive growth

(Supplementary Table 2), and additionally in one of the models (DSSAT) by an indirect effect of temperature on rate of pod addition at above-optimal temperatures (equivalent to fruit abortion). However, a primary factor contributing to yield reductions with increasing temperature was the shortening of crop life cycle duration (Supplementary Figs. 4a to 4e). Of interest, a few models showed a negative quadratic response to increasing mean growing season temperature that caused prolonged life cycles under extremely high temperatures (Supplementary Figs. 4a to 4e) that have been reported in controlled environment studies (Allen et al., 2018; Tacarindua et al., 2013; Thomas et al., 2010). This occurs because increase of mean daily temperature (~beyond 31 °C) slows progression of pod-set, seed development, and crop maturation. Lastly, an increasing temperature that shortened the life cycle caused reductions in maximum LAI below critical values in several models that would partially explain the decrease in biomass (Supplementary Figs. 5a to 5e).

Prior studies found that the main factors driving the large uncertainties in yield simulations under climate change were the different temperature responses for processes within the models (Asseng et al., 2013; Wang et al., 2017). The soybean models showed a large variation in the complexity and number of processes affected by temperature (Supplementary Tables 1 and 2). More mechanistic approaches in some models may also pose additional challenges to parametrize and evaluate processes with observed data. It is important to note that we evaluated yield responses to temperature under well-watered conditions to avoid a confounding effect of water stress, thus we would expect greater variability in model responses under drought stress. While our study evaluated the uncertainty in simulated yield responses under changing temperature and [CO$_2$] the accuracy of the response curves can only be validated against experimental data. Detailed model validation and parameterization with elevated temperature experiments under a wide

range of natural air temperatures or heated plots on the same genotype is needed to increase the validity of model-simulated temperature responses.

We observed a calibration effect on the shape of the yield-temperature response (See example in Fig. 7) that was unexpected and in contrast with previous maize studies where calibration had a limited effect on the yield temperature response (Bassu et al., 2014). However, it should be noted that the maize team (Bassu et al., 2014) tested the calibration effect on the model ensemble and not on individual models, while we analyzed calibration effect on individual models. In our study, as in the maize study, cultivar crop coefficients were modified as needed during *Full* calibration to match the observed LAI, biomass, and pod growth. Most models showed a decrease in the nRMSE for simulation of biomass and LAI after calibration (Supplementary Fig. 1). Particular to soybean models and our study, modelers also modified phenology when needed to match observed pod growth. We found that changes in phenology or leaf area development during calibration also modified the maximum LAI and the growing season duration in response to temperature (Supplementary Fig. 4 and 5), which may be partially associated to changes in the yield response to temperature after calibration. The effect of calibration on the yield response to temperature has not been explored in-depth previously, and could have implications on climate change predictions using crop models. This issue needs further investigation.

The model ensemble was useful to reduce uncertainty associated with individual models and calibration. The model ensemble simulated yield reductions of 0.1–5.3% per 1 °C temperature increase that followed the responses found in experimental studies with similar baseline temperatures (Fig. 6b). This range is also consistent with a U.S. climate assessment review by Hatfield et al. (2011), who concluded that soybean yield in the Southern USA (with mean baseline temperature of 26.7 °C) would be reduced by 2.4% for a 0.8 °C temperature increase, or 3.0% per 1 °C. We found several soybean temperature response experiments with yield data that can be compared with the model responses in our work (Fig. 6b). Of interest, some of those experiments showed high variability and greater yield reductions in response to temperature than those simulated in our work (see responses above 24 °C baseline in Fig. 6b), which highlights the need to evaluate soybean models under elevated temperature conditions. There are limitations when comparing simulated model ensemble response with field experimental data due to differences in soil, weather, and crop management. However, this comparison was still helpful to evaluate the temperature and [CO$_2$] responses in models. Further details regarding the field experimental data derived from the literature review are provided in supplementary section SI4 for elevated temperature experiments and [CO$_2$] experiments.

4.3. Response to [CO$_2$]

Eight soybean models had an asymptotic positive yield response to [CO$_2$], in line with extensive literature of [CO$_2$] effects on photosynthesis and yield, except for two models that had a linear response. Of interest, there were notable differences in the magnitude of the yield response to [CO$_2$] among the ten models (Fig. 8a). The [CO$_2$] increment from 360 ppm to 540 ppm is similar to the expected increase in [CO$_2$] levels by 2100 (360 ppm in 1995, 411 ppm at present (NOAA-ESRL, 2020), ~540 ppm lower limit of CO$_2$-equivalent under RCP 4.5 climate scenario in 2100 (IPCC, 2014)). Crop models differed greatly in the magnitude of the [CO$_2$] response (6–31% among models, averaged across sites, Fig. 8a). The presence of these large differences indicates inadequate approaches or parameterization of model responses to increase in [CO$_2$] and insufficient evaluation of some models with observed experimental data under elevated [CO$_2$] in soybean. In contrast to the results obtained for the temperature response, the calibration effect (*Blind* to *Full*) on the yield response to [CO$_2$] was relatively small to non-existent depending on crop model. The effect of calibration on the yield response to [CO$_2$] was an indirect result from the

optimization of cultivar traits that determine the rate and duration of leaf area development to mimic observed LAI and biomass over time under ambient [CO$_2$] conditions. This limited effect of calibration on the yield response to [CO$_2$] was expected and agrees with previous multi-model studies for other crops (Bassu et al., 2014). Prior studies have argued that models used for climate change predictions should require parametrization of a photosynthetic response to [CO$_2$] that is not cultivar-specific (Bassu et al., 2014).

There is a small proposed temperature by [CO$_2$] interaction on the rate of photosynthesis, based on the theory of photosynthetic enzyme function (Kirschbaum, 1994; Morison and Lawlor, 1999). Our study indicated that many models showed a greater yield response to [CO$_2$] with increasing mean growing season temperature (Fig. 9a). However, the extent of site effect was very strong for two models (LINTUL and MONICA), and modest for most other models (Fig. 9a). Differences in [CO$_2$] responses across models and locations cannot be attributed to variable levels of increased water use efficiency under high [CO$_2$], given that simulations were conducted under automatic irrigation to avoid the confounding effect of water stress.

The model ensemble was valuable to minimize uncertainty due to highly variable responses to [CO$_2$] of individual models. The ensemble simulated a yield increase by 18–19% that is similar to the response reported in a meta-analysis by Ainsworth et al. (2002) with [CO$_2$] increases from ambient (330–415 ppm) to elevated levels (540–600 ppm) (Fig. 8b). Some FACE studies (Bishop et al., 2015; Morgan et al., 2005; Ruiz-Vera et al., 2013) have reported a relatively lower yield increase compared to that simulated by the model ensemble, which may be partially attributed to water stress in some studies (Fig. 8b, Fig. 9b). This hypothesis is supported when comparing with responses obtained during wet seasons in the same FACE facility in Champaign, IL (Gray et al., 2016) (Fig. 8b). In addition, a recent review by Allen et al. (2020) provides evidence that the rapid, widely fluctuating [CO$_2$] in FACE experiments results in an under-estimation of the yield [CO$_2$] response compared to open-top and controlled-environment chambers by a ratio of 0.68. Thus, the 13% increase in soybean yield reported in a review of FACE studies by Ainsworth and Long (2020) with raising [CO$_2$] from 370 to 390–550 ppm would be 19% when corrected by a 0.68 ratio, similar to the yield increase obtained by the ensemble model in this study. The comparison of the ensemble model response to [CO$_2$] levels above 600 ppm with experimental data is limited to growth chambers studies (Baker et al., 1989; Xu et al., 2016) (Fig. 8b). Of interest, the ensemble-mean simulated responses to rising [CO$_2$] are similar for biomass and yield, which is inconsistent with literature showing reductions in harvest index under rising [CO$_2$] (Fig. 8b, Supplementary Table 8) and this requires further investigation. The FACE experiments (Bishop et al., 2015; Ruiz-Vera et al., 2013), and controlled environment studies in soil-plant-atmosphere research (SPAR) chambers (Ruiz-Vera et al., 2013; Xu et al., 2016) have shown relatively smaller yield increases in response to [CO$_2$] compared to increases in total aboveground biomass (Fig. 8b). Similarly, in a meta-analysis of FACE studies, Ainsworth et al. (2002) reported a seed yield increase of 24% and a decline in harvest index by 11% under optimal growth conditions (Fig. 8b). Overall, there is consistent research evidence that harvest index decreases in soybean under elevated [CO$_2$], whereas model-simulated harvest index was almost unaffected by [CO$_2$] changes. Further research is needed to properly understand the mechanism that would allow a reduction in harvest index under elevated [CO$_2$] in soybean crop models.

In summary, yield simulations from a model ensemble can reduce uncertainty in [CO$_2$] responses for climate change impact assessment; however, some models showed high deviations from the ensemble model response and the literature and would benefit from further development and evaluation with experimental data under elevated [CO$_2$]. Models were markedly different in their responses to temperature and [CO$_2$], and we found that the multi-model ensemble was useful for avoiding extreme responses from some models, unexpected effects from

calibration, and provided yield responses close to those reported in the literature for locations of similar baseline temperatures.

5. Conclusions

We conducted the first AgMIP soybean model inter-comparison with ten models and data from five contrasting environments. These models varied considerably in their yield responses to temperature and [CO_2] regardless of the level of calibration. While calibration generally did not affect temperature and [CO_2] response, it altered the shape of temperature response curves for some models and sites. The variability among models in their yield response to key climate change factors suggests that the individual models may currently have limited applicability for climate change food projections. Our study demonstrates that multi-model ensembles are valuable in reducing uncertainty in soybean yield simulations that may be associated with model and calibration. We also found that the temperature and [CO_2] responses of the model ensemble were close to those reported in the literature with some noted exceptions. Model ensembles, therefore, appear better suited than individual models in estimating climate change impacts on yield. Further improvement of soybean models by direct comparison with data from elevated [CO_2] and temperature experiments is needed to help parametrize and verify validity of responses to temperature and [CO_2] and to increase confidence in yield predictions under climate change.

CRediT authorship contribution statement

Kritika Kothari: Conceptualization, Formal analysis, Methodology, Visualization, Writing – original draft; **Rafael Battisti:** Conceptualization, Methodology, Software, Writing – original draft; **Kenneth J. Boote**: Conceptualization, Methodology, Funding acquisition, Writing – original draft; **Sotirios V. Archontoulis**: Investigation, Software, Writing – review & editing; **Adriana Confalone**: Investigation, Writing – review & editing; **Julie Constantin**: Software, Writing – review & editing, **Santiago V. Cuadra**: Software, Writing – review & editing, **Philippe Debaeke**: Software, Writing – review & editing; **Babacar Faye**: Software, Writing – review & editing; **Brian Grant**: Software, Writing – review & editing, **Gerrit Hoogenboom**: Software, Writing – review & editing; **Qi Jing**: Software, Writing – review & editing; **Michael van der Laan**: Software, Writing – review & editing; **Fernando Antònio Macena da Silva**: Investigation, Writing – review & editing; **Fabio R. Marin**: Software; Writing – Review & Editing; **Alireza Nehbandanim**: Software, Writing – Review & Editing; **Claas Nendel**: Software, Writing – Review & Editing; **Larry C. Purcell**: Investigation, Writing – review & editing; **Budong Qian**: Software, Writing – review & editing; **Alex C. Ruane**: Methodology, Writing – review & editing; **Céline Schoving**: Software, Writing – review & editing; **Evandro H.F. M. Silva**: Software, Writing – review & editing; **Ward Smith**: Software, Writing – review & editing, **Afshin Soltani**: Software, Writing – review & editing; **Amit Srivastava**: Software, Writing – review & editing; **Nilson A. Vieira Jr.**: Software, Writing – review & editing; **Stacey Slone**: Methodology, Formal analysis; **Montserrat Salmerón**: Conceptualization, Formal analysis; Funding acquisition, Investigation, Supervision, Methodology, Writing-original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.eja.2022.126482.

References

Ainsworth, E.A., Davey, P.A., Bernacchi, C.J., Dermody, O.C., Heaton, E.A., Moore, D.J., Morgan, P.B., Naidu, S.L., Yoo Ra, H.-S., Zhu, X.-G., Curtis, P.S., Long, S.P., 2002. A meta-analysis of elevated [CO_2] effects on soybean (Glycine max) physiology, growth and yield. Glob. Change Biol. 8 (8), 695–709. https://doi.org/10.1046/j.1365-2486.2002.00498.x.

Ainsworth, E.A., Long, S.P., 2020. 30 years of free-air carbon dioxide enrichment (FACE): what have we learned about future crop productivity and its potential for adaptation? Glob. Change Biol. 27 (1), 27–49. https://doi.org/10.1111/gcb.15375.

Allen, L.H., Kimball, B.A., Bunce, J.A., Yoshimoto, M., Harazono, Y., Baker, J.T., Boote, K.J., White, J.W., 2020. Fluctuations of CO_2 in Free-Air CO_2 enrichment (FACE) depress plant photosynthesis, growth, and yield. Agric. For. Meteorol. 284, 107899 https://doi.org/10.1016/j.agrformet.2020.107899.

Allen, L.H., Zhang, L., Boote, K.J., Hauser, B.A., 2018. Elevated temperature intensity, timing, and duration of exposure affect soybean internode elongation, mainstem node number, and pod number per plant. Crop J. 6 (2), 148–161. https://doi.org/10.1016/j.cj.2017.10.005.

Alsajri, F.A., Wijewardana, C., Irby, J.T., Bellaloui, N., Krutz, L.J., Golden, B., Gao, W., Reddy, K.R., 2020. Developing functional relationships between temperature and soybean yield and seed quality. Agron. J. 112 (1), 194–204. https://doi.org/10.1002/agj2.20034.

Archontoulis, S.V., Miguez, F.E., Moore, K.J., 2014. A methodology and an optimization tool to calibrate phenology of short-day species included in the APSIM PLANT model: Application to soybean. Environ. Model. Softw. 62, 465–477. https://doi.org/10.1016/j.envsoft.2014.04.009.

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P.J., Rötter, R.P., Cammarano, D., Brisson, N., Basso, B., Martre, P., Aggarwal, P.K., Angulo, C., Bertuzzi, P., Biernath, C., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Müller, C., Naresh Kumar, S., Nendel, C., O'Leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M.A., Shcherbak, I., Steduto, P., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J.W., Williams, J.R., Wolf, J., 2013. Uncertainty in simulating wheat yields under climate change. Nat. Clim. Change 3 (9), 827–832. https://doi.org/10.1038/nclimate1916.

Baek, J.-K., Sang, W.-G., Kim, J.-H., Shin, P., Cho, J.-I., Seo, M.-C., 2020. Yield response of soybean [Glycine max (L.) Merrill] to high temperature condition in a temperature gradient chamber. Korean J. Crop Sci. 65 (4), 339–345. https://doi.org/10.7740/kjcs.2020.65.4.339.

Baker, J.T., Allen Jr, L.H., Boote, K.J., Jones, P., Jones, J.W., 1989. Response of soybean to air temperature and carbon dioxide concentration. Crop Sci. 29 (1), 98–105. https://doi.org/10.2135/cropsci1989.0011183X002900010024x.

Bassu, S., Brisson, N., Durand, J.-L., Boote, K., Lizaso, J., Jones, J.W., Rosenzweig, C., Ruane, A.C., Adam, M., Baron, C., Basso, B., Biernath, C., Boogaard, H., Conijn, S., Corbeels, M., Deryng, D., De Sanctis, G., Gayler, S., Grassini, P., Hatfield, J., Hoek, S., Izaurralde, C., Jongschaap, R., Kemanian, A.R., Kersebaum, K.C., Kim, S.-H., Kumar, N.S., Makowski, D., Müller, C., Nendel, C., Priesack, E., Pravia, M.V., Sau, F., Shcherbak, I., Tao, F., Teixeira, E., Timlin, D., Waha, K., 2014. How do various maize crop models vary in their responses to climate change factors? Glob. Change Biol. 20 (7), 2301–2320. https://doi.org/10.1111/gcb.12520.

Battisti, R., Sentelhas, P.C., Boote, K.J., 2017. Inter-comparison of performance of soybean crop simulation models and their ensemble in southern Brazil. Field Crops Res. 200, 28–37. https://doi.org/10.1016/j.fcr.2016.10.004.

Battisti, R., Sentelhas, P.C., Boote, K.J., 2018. Sensitivity and requirement of improvements of four soybean crop simulation models for climate change studies in Southern Brazil. Int. J. Biometeorol. 62 (5), 823–832. https://doi.org/10.1007/s00484-017-1483-1.

Bishop, K.A., Betzelberger, A.M., Long, S.P., Ainsworth, E.A., 2015. Is there potential to adapt soybean (Glycine max Merr.) to future [CO_2]? An analysis of the yield response of 18 genotypes in free-air CO_2 enrichment. Plant, Cell Environ. 38 (9), 1765–1774. https://doi.org/10.1111/pce.12443.

Boote, K.J., Allen, L.H., Prasad, P.V.V., Baker, J.T., Gesch, R.W., Snyder, A.M., Pan, D., Thomas, J.M.G., 2005. Elevated temperature and CO_2 impacts on pollination, reproductive growth, and yield of several globally important crops. J. Agric. Meteorol. 60 (5), 469–474. https://doi.org/10.2480/agrmet.469.

Boote, K.J., Hoogenboom, G., Jones, J.W., Ingram, K.T., 2009. Modeling nitrogen fixation and its relationship to nitrogen uptake in the CROPGRO model. In: Ma, L.,

Ahuja, L.R., Bruulsema, T. (Eds.), Quantifying and understanding plant nitrogen uptake for systems modeling. CRC Press, Boca Raton, FL, pp. 13–46.

Burkey, K., Tisdale, R., Zobel, R., Ray, S., Pursley, W., 2020. Interactive effects of elevated ozone and temperature on growth and yield of soybean (Glycine max (L.) Merr.) under field conditions. Agronomy 10 (11), 1803. https://doi.org/10.3390/agronomy10111803.

Falconnier, G.N., Corbeels, M., Boote, K.J., Affholder, F., Adam, M., MacCarthy, D.S., Ruane, A.C., Nendel, C., Whitbread, A.M., Justes, É., Ahuja, L.R., Akinseye, F.M., Alou, I.N., Amouzou, K.A., Anapalli, S.S., Baron, C., Basso, B., Baudron, F., Bertuzzi, P., Challinor, A.J., Chen, Y., Deryng, D., Elsayed, M.L., Faye, B., Gaiser, T., Galdos, M., Gayler, S., Gerardeaux, E., Giner, M., Grant, B., Hoogenboom, G., Ibrahim, E.S., Kamali, B., Kersebaum, K.C., Kim, S.-H., van der Laan, M., Leroux, L., Lizaso, J.I., Maestrini, B., Meier, E.A., Mequanint, F., Ndoli, A., Porter, C.H., Priesack, E., Ripoche, D., Sida, T.S., Singh, U., Smith, W.N., Srivastava, A., Sinha, S., Tao, F., Thorburn, P.J., Timlin, D., Traore, B., Twine, T., Webber, H., 2020. Modelling climate change impacts on maize yields under low nitrogen input conditions in sub-Saharan Africa. Glob. Change Biol. 26 (10), 5942–5964. https://doi.org/10.1111/gcb.15261.

Fleisher, D.H., Condori, B., Quiroz, R., Alva, A., Asseng, S., Barreda, C., Bindi, M., Boote, K.J., Ferrise, R., Franke, A.C., Govindakrishnan, P.M., Harahagazwe, D., Hoogenboom, G., Naresh Kumar, S., Merante, P., Nendel, C., Olesen, J.E., Parker, P. S., Raes, D., Raymundo, R., Ruane, A.C., Stockle, C., Supit, I., Vanuytrecht, E., Wolf, J., Woli, P., 2017. A potato model intercomparison across varying climates and productivity levels. Glob. Change Biol. 23 (3), 1258–1281. https://doi.org/10.1111/gcb.13411.

Gaydon, D.S., Balwinder, S., Wang, E., Poulton, P.L., Ahmad, B., Ahmed, F., Akhter, S., Ali, I., Amarasingha, R., Chaki, A.K., Chen, C., Choudhury, B.U., Darai, R., Das, A., Hochman, Z., Horan, H., Hosang, E.Y., Kumar, P.V., Khan, A.S.M.M.R., Laing, A.M., Liu, L., Malaviachichi, M.A.P.W.K., Mohapatra, K.P., Muttaleb, M.A., Power, B., Radanielson, A.M., Rai, G.S., Rashid, M.H., Rathanayake, W.M.U.K., Sarker, M.M.R., Sena, D.R., Shamim, M., Subash, N., Suriadi, A., Suriyagoda, L.D.B., Wang, G., Wang, J., Yadav, R.K., Roth, C.H., 2017. Evaluation of the APSIM model in cropping systems of Asia. Field Crops Res. 204, 52–75. https://doi.org/10.1016/j.fcr.2016.12.015.

Grassini, P., Cafaro La Menza, N., Rattalino Edreira, J.I., Monzón, J.P., Tenorio, F.A., Specht, J.E., 2021. Chapter 8 - Soybean. In: Sadras, V.O., Calderini, D.F. (Eds.), Crop physiology case histories for major crops. Academic Press, London, UK, pp. 282–319.

Gray, S.B., Dermody, O., Klein, S.P., Locke, A.M., McGrath, J.M., Paul, R.E., Rosenthal, D. M., Ruiz-Vera, U.M., Siebers, M.H., Strellner, R., Ainsworth, E.A., Bernacchi, C.J., Long, S.P., Ort, D.R., Leakey, A.D.B., 2016. Intensifying drought eliminates the expected benefits of elevated carbon dioxide for soybean. Nat. Plants 2 (9), 16132. https://doi.org/10.1038/nplants.2016.132.

Greenland, S., 2019. Valid p-values behave exactly as they should: some misleading criticisms of p-values and their resolution with s-values. Am. Stat., 73(sup1) 106–114. https://doi.org/10.1080/00031305.2018.1529625.

Hao, X., Gao, J., Han, X., Ma, Z., Merchant, A., Ju, H., Li, P., Yang, W., Gao, Z., Lin, E., 2014. Effects of open-air elevated atmospheric CO2 concentration on yield quality of soybean (Glycine max (L.) Merr. Agric. Ecosyst. Environ. 192, 80–84. https://doi.org/10.1016/j.agee.2014.04.002.

Hatfield, J.L., Boote, K.J., Kimball, B.A., Ziska, L.H., Izaurralde, R.C., Ort, D., Thomson, A.M., Wolfe, D., 2011. Climate impacts on agriculture: implications for crop production. Agron. J. 103 (2), 351–370. https://doi.org/10.2134/agronj2010.0303.

IPCC. (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change: IPCC.

Jägermeyr, 2021. Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. Nature Food 873–885. https://doi.org/10.1038/s43016-021-00400-y.

Jeffers, D.L., Shibles, R.M., 1969. Some effects of leaf area, solar radiation, air temperature, and variety on net photosynthesis in field-grown soybeans. Crop Sci. 9 (6), 762–764. https://doi.org/10.2135/cropsci1969.0011183X000900060027x.

Kirschbaum, M.U.F., 1994. The sensitivity of C3 photosynthesis to increasing CO2 concentration: a theoretical analysis of its dependence on temperature and background CO2 concentration. Plant, Cell Environ. 17 (6), 747–754. https://doi.org/10.1111/j.1365-3040.1994.tb00167.x.

Li, T., Hasegawa, T., Yin, X., Zhu, Y., Boote, K., Adam, M., Bregaglio, S., Buis, S., Confalonieri, R., Fumoto, T., Gaydon, D., Marcaida Iii, M., Nakagawa, H., Oriol, P., Ruane, A.C., Ruget, F., Singh, B., Singh, U., Tang, L., Tao, F., Wilkens, P., Yoshida, H., Zhang, Z., Bouman, B., 2015. Uncertainties in predicting rice yield by current crop models under a wide range of climatic conditions. Glob. Change Biol. 21 (3), 1328–1341. https://doi.org/10.1111/gcb.12758.

Marin, F.R., Thorburn, P.J., Nassif, D.S.P., Costa, L.G., 2015. Sugarcane model intercomparison: structural differences and uncertainties under current and potential future climates. Environ. Model. Softw. 72, 372–386. https://doi.org/10.1016/j.envsoft.2015.02.019.

Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J.W., Rötter, R.P., Boote, K.J., Ruane, A.C., Thorburn, P.J., Cammarano, D., Hatfield, J.L., Rosenzweig, C., Aggarwal, P.K., Angulo, C., Basso, B., Bertuzzi, P., Biernath, C., Brisson, N., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R.F., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Müller, C., Nendel, C., O'Leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M.A., Shcherbak, I., Steduto, P., Stöckle, C.O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., White, J.W., Wolf, J., 2015. Multimodel ensembles of wheat growth: many models are better than one. Glob. Change Biol. 21 (2), 911–925. https://doi.org/10.1111/gcb.12768.

Mavromatis, T., Boote, K.J., Jones, J.W., Wilkerson, G.G., Hoogenboom, G., 2002. Repeatability of model genetic coefficients derived from soybean performance trials across different states. Crop Sci. 42 (1), 76–89. https://doi.org/10.2135/cropsci2002.7600.

Morgan, P.B., Bollero, G.A., Nelson, R.L., Dohleman, F.G., Long, S.P., 2005. Smaller than predicted increase in aboveground net primary production and yield of field-grown soybean under fully open-air [CO2] elevation. Glob. Change Biol. 11 (10), 1856–1865. https://doi.org/10.1111/j.1365-2486.2005.001017.x.

Morison, J.I.L., Lawlor, D.W., 1999. Interactions between increasing CO2 concentration and temperature on plant growth. Plant, Cell Environ. 22 (6), 659–682. https://doi.org/10.1046/j.1365-3040.1999.00443.x.

Moss, R.H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K., van Vuuren, D.P., Carter, T.R., Emori, S., Kainuma, M., Kram, T., Meehl, G.A., Mitchell, J.F.B., Nakicenovic, N., Riahi, K., Smith, S.J., Stouffer, R.J., Thomson, A.M., Weyant, J.P., Wilbanks, T.J., 2010. The next generation of scenarios for climate change research and assessment. Nature 463 (7282), 747–756. https://doi.org/10.1038/nature08823.

Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Folberth, C., Glotter, M., Hoek, S., Iizumi, T., Izaurralde, R.C., Jones, C., Khabarov, N., Lawrence, P., Liu, W., Olin, S., Pugh, T.A.M., Ray, D.K., Reddy, A., Rosenzweig, C., Ruane, A.C., Sakurai, G., Schmid, E., Skalsky, R., Song, C.X., Wang, X., de Wit, A., Yang, H., 2017. Global gridded crop model evaluation: benchmarking, skills, deficiencies and implications. Geosci. Model Dev. 10 (4), 1403–1422. https://doi.org/10.5194/gmd-10-1403-2017.

Müller, C., Franke, J., Jägermeyr, J., Ruane, A.C., Elliott, J., Moyer, E., Heinke, J., Falloon, P.D., Folberth, C., Francois, L., Hank, T., Izaurralde, R.C., Jacquemin, I., Liu, W., Olin, S., Pugh, T.A.M., Williams, K., Zabel, F., 2021. Exploring uncertainties in global crop yield projections in a large ensemble of crop models and CMIP5 and CMIP6 climate scenarios. Environ. Res. Lett. 16 (3), 034040 https://doi.org/10.1088/1748-9326/abd8fc.

NOAA-ESRL. (2020). National Oceanic & Atmospheric Administration, Earth System Research Laboratory: Mauna Loa CO2 annual mean data by Dr. Pieter Tans, NOAA/ESRL (www.esrl.noaa.gov/gmd/ccgg/trends/) and Dr. Ralph KEELING, Scripps Institution of Oceanography (scrippsco2.ucsd.edu/).

Purcell, L.C., Serraj, R., Sinclair, T.R., De, A., 2004. Soybean N2 fixation estimates, ureide concentration, and yield responses to drought. Crop Sci. 44 (2), 484–492. https://doi.org/10.2135/cropsci2004.4840.

Qian, B., Jing, Q., Smith, W., Grant, B., Cannon, A.J., Zhang, X., 2020. Quantifying the uncertainty introduced by internal climate variability in projections of Canadian crop production. Environ. Res. Lett. 15 (7), 074032 https://doi.org/10.1088/1748-9326/ab88fc.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A.C., Müller, C., Arneth, A., Boote, K.J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T.A.M., Schmid, E., Stehfest, E., Yang, H., Jones, J.W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proc. Natl. Acad. Sci. 111 (9), 3268. https://doi.org/10.1073/pnas.1222463110.

Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P., Antle, J.M., Nelson, G.C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G., Winter, J.M., 2013. The agricultural model intercomparison and improvement project (AgMIP): protocols and pilot studies. Agric. For. Meteorol. 170, 166–182. https://doi.org/10.1016/j.agrformet.2012.09.011.

Ruane, A.C., Goldberg, R., Chryssanthacopoulos, J., 2015. Climate forcing datasets for agricultural modeling: merged products for gap-filling and historical climate series estimation. Agric. For. Meteorol. 200, 233–248. https://doi.org/10.1016/j.agrformet.2014.09.016.

Ruane, A.C., Hudson, N.I., Asseng, S., Cammarano, D., Ewert, F., Martre, P., Boote, K.J., Thorburn, P.J., Aggarwal, P.K., Angulo, C., Basso, B., Bertuzzi, P., Biernath, C., Brisson, N., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R.F., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Kumar, S.N., Müller, C., Nendel, C., O'Leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Rötter, R.P., Semenov, M.A., Shcherbak, I., Steduto, P., Stöckle, C.O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J.W., Wolf, J., 2016. Multi-wheat-model ensemble responses to interannual climate variability. Environ. Model. Softw. 81, 86–101. https://doi.org/10.1016/j.envsoft.2016.03.008.

Ruíz-Nogueira, B., Boote, K.J., Sau, F., 2001. Calibration and use of CROPGRO-soybean model for improving soybean management under rainfed conditions. Agric. Syst. 68 (2), 151–173. https://doi.org/10.1016/S0308-521X(01)00008-7.

Ruiz-Vera, U.M., Siebers, M., Gray, S.B., Drag, D.W., Rosenthal, D.M., Kimball, B.A., Ort, D.R., Bernacchi, C.J., 2013. Global warming can negate the expected CO2 stimulation in photosynthesis and productivity for soybean grown in the Midwestern United States. Plant Physiol. 162 (1), 410. https://doi.org/10.1104/pp.112.211938.

Sall, J. (2002). Monte Carlo calibration of distributions of partition statistics. Retrieved from (https://www.jmp.com/content/dam/jmp/documents/en/white-papers/montecarlocal.pdf).

Salmerón, M., Purcell, L.C., Vories, E.D., Shannon, G., 2017. Simulation of genotype-by-environment interactions on irrigated soybean yields in the U.S. Midsouth. Agric. Syst. 150, 120–129. https://doi.org/10.1016/j.agsy.2016.10.008.

Setiyono, T.D., Cassman, K.G., Specht, J.E., Dobermann, A., Weiss, A., Yang, H., Conley, S.P., Robinson, A.P., Pedersen, P., De Bruin, J.L., 2010. Simulation of soybean growth and yield in near-optimal growth conditions. Field Crops Res. 119 (1), 161–174. https://doi.org/10.1016/j.fcr.2010.07.007.

Storck, L., Cargnelutti Filho, A., Lúcio, A.D.C., Missio, E.L., Rubin, S.D.A.L., 2010. Avaliação da precisão experimental em ensaios de competição de cultivares de soja. Ciência e Agrotecnologia 34, 572–578.

Tacarindua, C.R.P., Shiraiwa, T., Homma, K., Kumagai, E., Sameshima, R., 2013. The effects of increased temperature on crop growth and yield of soybean grown in a temperature gradient chamber. Field Crops Res. 154, 74–81. https://doi.org/10.1016/j.fcr.2013.07.021.

Team, R.C., 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (URL). ⟨https://www.R-project.org/⟩.

Thomas, J., Boote, K., Pan, D., Allen Jr, L., 2010. Elevated temperature delays onset of reproductive growth and reduces seed growth rate of soybean. J. AgroCrop Sci. 1, 19–32.

Wallach, D., Martre, P., Liu, B., Asseng, S., Ewert, F., Thorburn, P.J., van Ittersum, M., Aggarwal, P.K., Ahmed, M., Basso, B., Biernath, C., Cammarano, D., Challinor, A.J., De Sanctis, G., Dumont, B., Eyshi Rezaei, E., Fereres, E., Fitzgerald, G.J., Gao, Y., Garcia-Vila, M., Gayler, S., Girousse, C., Hoogenboom, G., Horan, H., Izaurralde, R. C., Jones, C.D., Kassie, B.T., Kersebaum, K.C., Klein, C., Koehler, A.-K., Maiorano, A., Minoli, S., Müller, C., Naresh Kumar, S., Nendel, C., O'Leary, G.J., Palosuo, T., Priesack, E., Ripoche, D., Rötter, R.P., Semenov, M.A., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Wolf, J., Zhang, Z., 2018. Multimodel ensembles improve predictions of crop–environment–management interactions. Glob. Change Biol. 24 (11), 5072–5083. https://doi.org/10.1111/gcb.14411.

Wang, A., Lam, S.K., Hao, X., Li, F.Y., Zong, Y., Wang, H., Li, P., 2018. Elevated CO_2 reduces the adverse effects of drought stress on a high-yielding soybean (Glycine max (L.) Merr.) cultivar by increasing water use efficiency. Plant Physiol. Biochem. 132, 660–665. https://doi.org/10.1016/j.plaphy.2018.10.016.

Wang, E., Martre, P., Zhao, Z., Ewert, F., Maiorano, A., Rötter, R.P., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Aggarwal, P. K., Anothai, J., Basso, B., Biernath, C., Cammarano, D., Challinor, A.J., De Sanctis, G., Doltra, J., Dumont, B., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A.-K., Liu, L., Müller, C., Naresh Kumar, S., Nendel, C., O'Leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ripoche, D., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P., Waha, K., Wallach, D., Wang, Z., Wolf, J., Zhu, Y., Asseng, S., 2017. The uncertainty of crop yield projections is reduced by improved temperature response functions. Nat. Plants 3 (8), 17102. https://doi.org/10.1038/nplants.2017.102.

Xu, G., Singh, S.K., Reddy, V.R., Barnaby, J.Y., Sicher, R.C., Li, T., 2016. Soybean grown under elevated CO_2 benefits more under low temperature than high temperature stress: varying response of photosynthetic limitations, leaf metabolites, growth, and seed yield. J. Plant Physiol. 205, 20–32. https://doi.org/10.1016/j.jplph.2016.08.003.