Biogeosciences Discuss.,
doi:10.5194/bg-2016-527-RC1, 2017

Biogeosciences
Discussions
Open Access

Interactive comment on "Uncertainties in global crop model frameworks: effects of cultivar distribution, crop management and soil handling on crop yield estimates" *by* Christian Folberth et al.

Anonymous Referee #1

Received and published: 21 February 2017

In this article, the authors consider 5 evolutions of the EPIC model with which they perform identical simulations for global maize growth under different setups. One setup aims at keeping each model's as it is originally set and used with harmonized climate data, two setups aim at reducing the differences between the models by 1/ harmonizing N and P application rates and planting dates, 2/ also providing sufficient nutrients to avoid any yield decrease. The simulations are also performed under rainfed and irrigated conditions. The study is centered on the coefficient of variation describing the spread of the models in the different configurations.

Printer-friendly version

Discussion paper

About the content, the study presented here might be of interest for the EPIC community but lacks scientific body for the wider biogeoscience and modeling communities. The scientific question is not clear and the focus of the paper is only on the variation of yield simulation between same-family models. This variation is found to be lower in simulations when there is no nutrient or water deficit, which is expected since inhibiting two crucial model processes involved in yield simulation. Parameters, run setups, cultivar distributions, are combined, tested, permutated, without seemingly thinking about their physical meaning, using the models as black boxes from which to turn buttons. The environmental variables such as temperature, precipitation, water and N deficits are omitted all along the analysis preventing a deep analysis of the causes of the biases or differences between models. This article focuses on trying to reduce the differences in yield estimates between same-family models without trying to understand the processes that might be under play and if one of the initial models might be more justified than the others for this type of study. Some of these background questions are touched in the discussion section but are not supported by any analysis or even literature review.

About the style, a number of inconsistencies and imprecisions could be noted. The literature review both in the introduction and discussion does not show a high level of analysis of state-of-the-art research and weakens the potential results of the study. Below is a detailed list of comments and suggestions for revision of the manuscript.

Introduction About the GGCMs in general L67 : 'For some models . . .' The reader is expecting a 'For others'. It should be clear from here that only one type of models will be mentioned here.

L71-77 : Another limitation of this type of model is ignored in this section, only to be a shortly mentioned in the discussion and that is the heterogeneity of agricultural practices. This study does not cover this aspect which is justified by its complexity that requires different methods, but taking this factor into account is crucial for analyzing the performance of global crop models. Even with better data coverage and compilation, the spatial resolution of global gridded crop model might not compatible with the spatial heterogeneity of agricultural plots and practices.

L81-86 : this section is not clear and needs to be rephrased to make clear 'what affects what' in the real world and how it is represented in the model at which cost. In particular 'To limit such effects' does not sound appropriate.

About crop models' uncertainties In the introduction, it is not clear if the authors have studied the state of the art of crop model uncertainty analyses even if it is the core of their work. It is necessary that their work be put into the context of recent work on the topic and references must be listed. A range of studies have looked at parameterization and configuration of crop models for large-scale areas. Examples include Izumi et al., Ag. For. Met., 2009, Osborne et al., Ag. For. Met., 2013, Vallade et al., GMD 2014, Wang et al., 2013, , Zhao et al., Ecol. Mod., 2014. This lack of literature analysis results in a weak description of the problematic addressed by the article.

Introduction & methods About the models used for the analysis and their descriptions In the beginning the authors describe GGCM as the combination of a field-scale crop model and a model framework (L68-69). This description would have an interest if the uncertainty analysis then studied separately the uncertainty due to each component but it does not seem to be the case since all models used in the study are listed as different in their MFW and field-scale crop models. Then using this denomination of MFW is actually confusing because it does not seem appropriate with the use that is done of the term. From the introduction to section 2.2 it looks like the models run in the study are actually not model frameworks but model configurations with only routines activated or not and parameter values changes. However, when reading each model description, it appears that they have different grids (not clear from the descriptions), different scales of applications from field-scale to global-scale, different crop cultivars, and the processes included are very different from one model to another leading to the impression that they are actually today different models, from a same family, and derived from a same initial model, but not only model frameworks that 'process input data and run the model' as defined at L69. Table 1 should be extended with the details on the grid and other differences between models such as processes included or not, input data required or not. Either the definition of a MFW needs to be revised or another term needs to be used in the text that is more consistent with the diversity of models used. Table 1 would be much improved if a separation would be added between subsections such as" subroutines used", "parameter values", "processes included", "forcing data","cultivars distribution". . .

Methods L150-153 : "contribution to at least three management setups". Not clear. Does this refer to Agmip or ISI-MIP exercices? L235 : harmonized runs have not been described yet. L289 : The word 'de-trended' does not seem appropriate or the method is not described correctly. If I am correct, a moving average can smoothen data, not de-trend it. The word de-trend is used several times in the manuscript. If not the correct word it needs to be corrected all along, if it is a de-trending that is done then the methods needs to be explained in mode details. L293 : the percentage of variance explained is r2 and not r as stated here. Hence, a threshold of r>0.5 actually means r2>0.25, hence 25% of the variance is explained. L294 : no expression is given for CV.

Results The trend seen in EPIC-BOKU and PEPIC is hard to see and not supported by statistical analysis. It does not seem significant and does not seem worth discussing. The decrease in the range of simulations with increased setup harmonization is expected since sources of variability are removed. However what is more revealing is that all models move away from the reported yields with harmonized values. See discussion.

L336 : "Spatially, the deviation of maize yield estimates among the MFWs is largest with the default setups in tropical and arid regions Âż & L: 342 : "where (a) yields are at moderate or high levels, (b) most models plant the same high-yielding cultivar and (c) the annual temperature curve usually defines a narrow growing season window. Âż These are not supported by the data shown on the map (which is hard to get information from at this scale). A scatterplot of CV in for example precipitation/temperature or yield/growing season, irrigation/fertilization rates spaces are needed to support such claim. “in the low fertilizer input region Western Russia EPIC-IIASA, which plants here a cultivar adapted to colder climates, provides high yields in comparison to the other EPIC-MFWs, which raises CV from around 30-40% in the default setups to often 50-60% in the fullharm ensemble..” Refer to the adequate figure, because again this is hardly supported by Figure 2.

‘The patterns observed in Figure 3 indicate that yield estimates differ especially in low-input regions.’ Not supported by Figure 3.

Figure 5 is unreadable with so many panels and no background grid. Some lines or rows should be merged together with colors added to have more lines plotted in less windows, hence making different setups possible to compare. For example, this claim : ‘The choice of soil data (SoilD; rows one and two vs. three and four) has an impact on inter-annual yield 400 dynamics, especially if the parameterization and soil handling of EPIC-IIASA (column one) is used’ is not obvious from the figure. L401: ‘less apparent’. This part of the analysis lacks numbers the so-called effect of soil data needs to be quantified if discussed. Same for L411:‘higher peaks’.

Presenting the performance of the models in terms of number of countries above a given threshold is not relevant to the global crop modeling issue. Total arable area for example would be more interesting or at least taking into account size of the country. The question raised with figure S3-2 is for example more interesting and raised the real question that should be behind this study which is can a single crop model with a single world-wide parameterization be better everywhere? This discussion

Discussion This paragraph points to the problematic of having a world-wide set of parameters and 5° grids to describe small-scale agriculture and points to the limitation of the study which only considers the models from the setup angle.

A deeper analysis could have looked at the nitrogen and water deficits and correlate it with the driver of model differences. L506 : The EPIC model in contrast was specifically developed to investigate impacts of crop management on yields and due to the same set of algorithms employed in the EPIC-MFWs an increasingly good fit could be expected with increasing level of harmonization. There is no reason why increasing harmonization leads to better performance if the harmonization is done with not appropriate data. That is exactly what you see on Figure 2 with more harmonization leads to obviously a reduction of the models' spread but not at all to a smaller distance from reporter data.

"low yields and tropical climate are characteristic of regions in which the EPIC-MFWs show large differences in maize yield estimates in their default setups (Figure 3).". Nowhere Figure 3 shows a link between yield intensity and differences between yield estimates. L555 : None of these assertions is supported by either literature or your own analysis. L559: Parameterization of each individual model should have been a key hypothesis of the study. Were the harmonized values picked adequately? There are almost no references in the discussion to support any claim.

Conclusion L669 : The differences in model outputs induced by differences in setups indicate that further steps of harmonization 670 among GGCMs should be taken if model algorithms are to be compared globally.

Why would harmonization allow comparison? Harmonization can allow to identify sources of uncertainty but if the goal was harmonization why developing so many models in the first place? Each model has strengths and weaknesses but each model has been developed for specific configurations. A model is often calibrated and developed with a set of input data and simply giving all models the same input data does not help in understanding what the models do nor in making them better. Instead of harmonization, improvement of data, with lower-scale, or subscale heterogeneity parameterizations or processes improvements would improve the ability of the community to model global crop yields.

Interactive comment on Biogeosciences Discuss., doi:10.5194/bg-2016-527, 2016.