Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2016-527-AC1, 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.

Christian Folberth et al.

folberth@iiasa.ac.at

Received and published: 9 June 2017

We thank the two anonymous reviewers for their thoughtful, critical comments and have revised the manuscript accordingly. As addressing the comments required rewriting most of the manuscript, except few parts of the Methods and Results sections, we did not track changes to allow for better readability. Instead, we point out herein where Comments have been addressed predominantly in the new version of the manuscript, which is provided as a supplement linked at the end of this response. Due to the substantial extension of analyses, we have removed the supplementary results for wheat

Printer-friendly version

Discussion paper

in order to keep the already long manuscript at a reasonable length. Below we provide a point-by-point response to the reviewers' comments and suggestions.

Anonymous Referee 1

Synopsis In this article, the authors consider 5 evolutions of the EPIC model with which they per- form 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.

General comment 1) 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.

We have now stated the objectives of the study clearly in a new Section 1.4. While the focus is still on differences between same-family models and effects of harmonization on their comparability, we have substantially expanded on background information (e.g. Sections 2.4 and Supplementary Information (SI) 1) and analyses of factors driving differences among the models (Sections 3.3 and 3.5) besides extending analyses to a wider range of global gridded crop models (GGCMs) presented in Section 3.6 and included in the Discussion. We are convinced that this will render the study relevant also for the wider biogeosciences and crop modelling communities. While the EPIC model presently has the most detailed representation of soil routines within the ensemble an increasing number of modeling groups is in the course of implementing such processes in their models, rendering the study also of relevance for such efforts.

General comment 2) 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 basic ensemble simulations (see Section 1.3) had been setup to the best knowledge of the participating research groups and hence do not follow a systematic permutation of parameters but reflect rather the wide ranging assumptions on optimal setups of GGCMs GCM research groups. We consider this an asset of the study compared to narrowly targeted sensitivity analyses of crop models typically performed in single model studies as is now described in the Introduction (Section 1.2.2). While the latter is vital for studying model behavior with respect to singled out subroutines, our approach reflects a substantially wider range of options considered when setting up a GGCM. We agree that the more detailed permutation of setup options between the two EPIC-based GGCMs EPIC-IIASA and GEPIC lacked a detailed description of considered parameters and subroutines and that parameters had been treated as a bulk. Also the relevance of the parameters had not been presented in sufficient detail. Accordingly, we have now repeated this experiment splitting parameters into nutrient/organic matter turnover related and hydrologic parameters and treated the soil handling more consistently (Section 2.7). More detailed background information on the parameterizations of each EPIC-based GGCM is provided in Section 2.4 and parameters and subroutines have now been grouped systematically in Table 2. Relevant routines are laid out in SI 1.1 to allow for an understanding of each parameter's relevance.

General comment 3) 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.

We agree that such diagnostic variables are relevant for understanding actual drivers in model disagreement. Fertilizer supply had been included before and is still part of the analyses (e.g. Section 3.3). Climate data are identical in all simulations and pose hence not a source of uncertainty. We include now also precipitation as an important variable driving inter-annual yield variability (e.g. Frieler et al., accepted) in the same section, but found limited impact here. Crop yield estimates are the only variable that had been submitted by all participating modeling groups, which limits the feasibility of more detailed analyses of diagnostic variables within the ensemble in this study. However, for the evaluation of the importance of setup domains of two EPIC-based GGCMs, we are now including plant growth stress outputs from the model to provide a more thorough understanding of what the different setups cause in environmental processes and finally yield estimates (Section 3.5.1). As the aggregated magnitudes of plant growth stresses at the global scale have only limited explanatory power (see Section 3.5.1), we are also providing selected examples at the grid cell level in the SI (Figure S3-12 to S3-14).

General comment 4) 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.

We are now stating more clearly in the new objectives section (Section 1.4) that the study aims at identifying the impact of assumptions and parameterizations on yield estimates in GGCMs in order to derive future pathways for further harmonization. Harmonization itself solely serves for reducing selected differences (growing seasons and fertilizer supply) in order to better quantify the impact of remaining differences in model routines related to e.g. soil handling and parameterization or cultivar distributions. This allows for the identification of priorities for future ensemble harmonization, which is required to better identify model bias introduced by actual plant growth processes. The literature review has been extended accordingly (Section 1.2). The Results section has been extended substantially and now also covers the impact of setup components on plant stresses and - in an exemplary way - environmental variables (Section 3.5.1). In addition, we have added a sensitivity analyses for the permutation of setup domains among two EPIC-based GGCMs to highlight the spatial disparity in optimal model setups (Section 3.5.2). The Discussion has been updated accordingly.

General comment 5) 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.

We regret earlier imprecisions and lack of background information and have now elaborated on both (see below).

Comment 6) 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.

Based on yours and the second reviewer's suggestions, we have now rewritten the manuscript in most parts. The prior sentence is not included in the new version.

Comment 7) 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.

We include now more detailed information on management assumptions in the model descriptions (Section 2.4; Table 2, Table S1-3). We agree that single GGCMs can presently not reproduce management heterogeneity at the sub grid level, which renders

Printer-friendly version

Discussion paper

the use of an ensemble with widely differing setups an asset. However, the setups of the GGCMs are aiming at representing certain management systems in general but at the same time reproduce a global picture, which renders setups not always fully consistent. This is stated now in the Discussion (end of Section 4.1): "In summary, the EPIC-based GGCMs can be considered contrasting representations of agricultural systems and agro-ecologic conditions globally. Due to lack of spatial parameterizations except for a limited number of cultivars, however, none of them can be expected to represent an optimal setup. Instead, they allow for bracketing uncertainties that may exist at small scales (Ewert et al., 2011), e.g. through the representation of low-input or high-input agricultural systems and associated plant growth limitations within a pixel, which GGCMs cannot consider in their presently too coarse spatial resolution. For a more targeted representation of agricultural systems, however, setups should be compiled in a more consistent way. E.g., GEPIC is rather representative for smallholder systems but the narrow planting densities and ploughing operations are more common to high-input agriculture."

Comment 8) 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.

We have now rewritten most parts of the manuscript and the prior statement in not included anymore.

Comment 9) 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

Printer-friendly version

Discussion paper

article.

We agree that the literature review had been very limited in the earlier version of the manuscript and thank the reviewer for the suggestions. We have now substantially expanded on this and included the suggested references in the Introduction (Section 1.2) and Discussion (Sections 4.5 and 4.6).

Comment 10) 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”. . .

We agree that the term “model framework” was not used appropriately in this context in the earlier version of the manuscript. In accordance with your suggestion, we are now considering each instance also of the EPIC-based models a unique GGCM and refer to them accordingly as “EPIC-based GGCMs”. We have also restructured the former Table 1 (now Table 2) into parameter and subroutine domains. Cultivars are still addressed separately (Section 2.5) to account for their spatial distributions and the use of multiple cultivars in most GGCMs. Table 3, which describes the setup domains permutated between EPIC-IIASA and GEPIC has been restructured as well to comply with the parameter definitions of the single EPIC-based GGCMs in Table 2. Default forcing data are still included in the GGCM rationales (SI 1.3) and the data section (Section 2.6)

Comment 11) Methods L150-153: “contribution to at least three management setups”. Not clear. Does this refer to Agmip or ISI-MIP exercices?

We have now restructured the description of the experiment with general background information on harmonization and selection of GGCMs for this study in a new Section 1.3, in which also the crop modelling initiatives the experiment is based on are briefly pointed out.

Comment 12) L235: harmonized runs have not been described yet.

See response to your comment 11.

Comment 13) 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.

Thank you for pointing this out. We state now correctly (Section 2.8.2): “Reported yields were de-trended by subtracting the 5-year moving mean in order to remove trends in yields due to changes in technology and management (Elliott et al., 2015;

Printer-friendly version

Discussion paper

Müller et al., 2017)."

Comment 14) 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

Thank you for pointing this ouot. Rather than using the explained variance and a self-defined threshold, we now use Pearson's correlation coefficient r directly and the significance threshold of $p<0.1$ as a benchmark for model performance following the model evaluation framework presented recently in Müller et al. (2017).

Comment 15) L294: no expression is given for CV.

The coefficient of variation (CV) is now included in the Methods section (Eq. 2; Section 2.8.1).

Comment 16) 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.

We have now added linear regressions, which are included in a supplementary figure (Figure S3-1) for better readability. Regression coefficients are presented in Table S3-2, which also includes the mean bias. The coefficients show that there are clear trends in some of the models depending on the management scenarios, indicating incompatibility of their default setups with harmonization which also applies to non-EPIC-based GGCMs (see Discussion, Section 4.3). It is correct that some of the models move away from reported yields after harmonization, foremost in the scenario with sufficient nutrient supply. However, the purpose of harmonization as such is not to achieve a better agreement with reported yields, but to eliminate differences among models. The mean bias from reported yields is hence foremost of relevance in the default setup and the harmonized setup with reported fertilizer application rates (fullharm). The now also included analyses of time-series correlation show also that model agreement concerning this metric increases if sufficient nutrients are supplied (see Section 3.2.2). In addition, the now extended results of the model benchmarking indicate that models taking stress interactions into account apparently perform better in reproducing reported inter-annual yield dynamics if sufficient nutrients are supplied (Section 3.4 and 3.6). This poses a dilemma in model evaluation, which may have to be dealt with as discussed now in a new Section 4.5.

Comment 17) L336 : "Spatially, the deviation of maize yield estimates among the MFWs is largest with the default setups in tropical and arid regions L342: "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." 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

We agree that the maps alone did not provide sufficient information to support these claims. The earlier shown plots of fertilizer application rates compared to CV include now linear regressions (Figure 5), which is also the case for a newly added comparison of time-series correlation among models (Figure 6). We are now also providing a wide range of supplementary information relating the findings to cultivar distributions, precipitation, and fertilizer supply in Figure S3-2 to S3-8 to support the earlier statements. In addition, fertilizer application rates of the harmonized input datasets are displayed in Figure S2-2a,b and a map of climate regions in Figure S2-1.

Comment 18) 'The patterns observed in Figure 3 indicate that yield estimates differ especially in low-input regions.' Not supported by Figure 3. We regret this inconsistency.

This was supposed to refer to earlier Figure 4, now Figure 5, which shows that the largest deviations occur at low fertilizer application rates. For the visual interpretation, maps of harmonized fertilizer application rates are now provided in Figure S2-2a,b.

Comment 19) 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 L400: ‘The choice of soil data (SoilD; rows one and two vs. three and four) has an impact on inter-annual yield 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’.

We agree that this section (now Section 3.5) did not provide sufficient background information. The analysis of setup domains has been extended by an additional dimension to separate parameters for nutrients/organic matter turnover and hydrology (see response to your comment 2 and Section 2.7). Due to the added complexity, we were not able to reduce the number of panels but added a more pronounced background grid to all figures to allow for better readability. In addition, we present now the relative deviation in yield estimates from the basic EPIC-IIASA setup to better visualize differences whereas absolute yields have been moved to the SI (Figure S3-10). The figure now also includes the mean bias (ME) from the reference simulation and a normalized coefficient of variation over time (CVt) as a metric for changes in inter-annual yield variability. Besides this quantification, we also produced a correlation matrix of global average yields among all considered setups to identify the most sensitive setup components (Figures S3-15 and S3-16). All sections relating to this topic have been rewritten accordingly.

Comment 20) 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?

We agree that the global picture as such did not provide much insight. Although we consider all countries relevant as GGCM outputs are typically interpreted globally, the quality of input and benchmark data is often poorest in low-input regions. We are hence now focusing on the major maize producing countries and have moved the suggested figure (now Figure 7) to the main manuscript. Other figures were adjusted to present benchmarking results accordingly at the same scale (Figure 8; Figure 10; Table S3-5; Figure S3-17; Figure S3-22). A wider range of countries for which reported yields have not been estimated by FAO and which did not experience a substantial change in harvested areas is included in the supplementary (Figure S3-9). The extended analyses highlight that no single model performs best everywhere, which can be expected from the lack of truly spatial parameterizations, and that nutrient supply has an impact on performance for models considering stress interactions as is the case for the EPIC-based but apparently also other GGCMs. This emphasizes on the one hand the advantage of using crop model ensembles for impact studies (see Discussion, e.g. Section 4.1), and on the other hand suggests that model evaluation may have to be carried out across different setups as suggested now in the discussion (Sections 4.4 and 4.5).

Comment 21) Discussion: This paragraph points to the problematic of having a world-wide set of parameters and 5âŮę 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.

We agree and have now extended the analysis of model differences between two selected EPIC-based GGCMs to plant growth stresses. This cannot be related to yield estimates directly, which also depend on the estimation of potential biomass and the timing of the stress (see Section 3.5.1) but at least to certain parameterizations. As this cannot be analyzed at the global scale in detail, we are providing examples at the grid level (Figures S3-12 to S3-14).

Comment 22) 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.

In the new version of the manuscript, the Discussion has been rewritten entirely and the specified statement is not include anymore. For the effect of harmonization, see also our reply to your comment 16.

Comment 23) “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.

This was supposed to refer to earlier Figure 4 (now Figure 5), which shows model ensemble bias compared to fertilizer application rates.

Comment 24) L555: None of these assertions is supported by either literature or your own analysis.

We have now rewritten the Discussion entirely and reference the findings presented in the Results and the literature more thoroughly.

Comment 25) 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.

As stated in the new objectives (Section 1.4), the experimental design of the study does not focus on identifying an optimal setup or its adequacy. In fact, the feasibility to do this globally is limited especially by having only reported yields as a benchmark variable, which are in addition of limited quality in various parts of the world (e.g. Section 2.8.2). We are now presenting (Section 2.4) and discussing (Section 4.1) the setups, their rationales, and advantages/disadvantages more thoroughly. While specific model parameterizations are typically a key hypothesis in single model studies, e.g. investigating model sensitivity to a priori selected parameters, we are focusing on a wide ensemble to dissect differences among models and their implications, and derive implications for future research in ensemble and single model studies. The harmonization is based on the presently most up-to-date global data on crop management, which should justify their adequacy. In general, however, we would argue that as long as harmonization data are within a reasonable range, they should be adequate to allow for eliminating selected aspects in differences of model setups, which was shown to be the case here.

Comment 26) Conclusion L669: The differences in model outputs induced by differences in setups indicate that further steps of harmonization 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.

As stated in our response to your last comment, harmonization allows for narrowing the range of uncertainties among models and hence for learning in a more targeted way, where differences arise from. We agree and show in the manuscript that harmonization can cause problems in combination with default model calibrations (e.g. Section 4.3). However, e.g. to study model response to drought conditions will certainly benefit from harmonizing soil data, which were shown here to have a substantial impact on water supply in the EPIC-based GGCMs. This can be also expected from the wider ensemble, in which almost all models use at least soil texture as an input. If these data are not harmonized, it will hardly be feasible to compare plant water stress representations in different models unless a wide range of diagnostic variables is supplied.

Anonymous Referee 2

Synopsis The poor presentation of “Uncertainties in global crop model frameworks: effects of cultivar distribution, crop management and soil handling on crop yield estimates” hampered me from making an in-depth evaluation of the research presented in the manuscript. The presentation of the study should be largely improved before this study becomes acceptable for publication in any outlet.

We have now rewritten the manuscript to a large extent and elaborated further on the presentation and extent of background information, analyses, and discussion.

General comment 1) From my current understanding of the study, the research itself would also require a lot of additional work to become of interest to the readership of Biogeosciences.

We regret to hear that the manuscript did not meet the reviewer’s expectations at all and that only limited suggestions on how to further improve it were supplied. We have hence focused on the suggestions of Reviewer 1, including your general comments, and extended the analyses to our best knowledge to carve out the relevance for the wider biogeosciences community.

General comment 2) The manuscript does not mention the objectives of the study. Furthermore, it was not possible to derive the objectives from the structure, the results or the discussion. Any revision should be structured around clear objective(s).

We agree that the objectives had not been stated in sufficient clarity before and have now added a brief section on this (Section 1.4) that builds on a review of recent research in the field (Section 1.2).

General comment 3) The analyses are very much focused on different versions of the EPIC model, it is not clear what the readership outside the EPIC community can learn from this analysis.

We have now expanded the sections on background information and experimental design (Sections 1 and 2) to better embed the study in on-going research. While the focus is still on five EPIC-based global gridded crop models (GGCMs), we have (a) extended the analyses to a wider ensemble of GGCMs based on different core models (Sections 2.1 and 3.6) and (b) deepened analyses of differences among EPIC-based GGCMs (Section 3.5) to highlight the relevance of subroutines and data that may apply also to other models including similar components. In fact, only the availability of five different EPIC-based GGCMs within the ensemble allows for studying the role of parameter choices made by individual modeling groups in comparison to choosing a specific core model, which is an aspect typically ignored in other model intercomparison studies and thus an asset of this work.

General comment 4) The study includes a couple of GGCM but it was not clear what these models contributed to the analysis as a whole. It appears as if they were added to satisfy administrative issues rather than their contribution to the science. We agree that the wider ensemble of GGCMs had been included in a very limited way in the earlier version of the manuscript and have extended the analyses, foremost in the new Section 3.6 and the Discussion (Section 4) with key characteristics of the wider ensemble presented in Table S1-1.

General comment 5) The analyses appears superficial; differences between models are reported but not explained, it is not clear which of the differences should be considered valuable (e.g., representing a lack of process understanding) and which

Printer-friendly version

Discussion paper

problematic (e.g., a poor representation of a well-known process). In the absence of clear objectives, the analysis itself appears flawed.

The objectives of the study are now specified in a new Section 1.4: "...to (a) identify key assumptions that drive differences in yield estimates in the EPIC-based GGCM ensemble, (b) derive priorities for further improvements in model input data (e.g. management aspects) or – in the absence of suitable data sets – harmonization in the wider GGCMI ensemble to address central uncertainties as a set of scenarios rather than random choices made by modelers (Confalonieri et al. 2016), and (c) provide a thorough documentation of GGCM setups and their implications as an aid in the potential further use of the publically available GGCMI phase 1 outputs and in the interpretation of earlier studies the participating models were used in (e.g. Elliott et al., 2014; Rosenzweig et al., 2014; Deryng et al., 2016; Müller et al., 2017; Schauberger et al., 2017; Frieler et al., forthcoming)." Identifying a lack of process understanding is not the purpose of this study. In contrast we are aiming to identify model processes and data that impact plant growth and yield estimates without being part of the plant growth and yield formation algorithms directly. This allows for evaluating the impact of model harmonization and for identifying further input data that may be of relevance for harmonization in order to compare model routines with regard to adequacy and differences in future studies.

General comment 6) Given that all of the EPIC models have a common ancestor, one could expect than harmonization will decrease the CV of the results because the harmonization mimics going back in time. If this approach would be used for models which are thought to be independent, such an exercise could re- veal that they have a common ancestor anyway (for example, the same photosynthesis model). Failure to substantially decrease the CV following harmonization could then be interpreted that the tested models are indeed independent. What can be learned by applying this method on models that are known to be dependent? In my opinion the discussion lacks a biogeochemical component.

The harmonization approach may not have been presented in sufficient detail in the earlier version of the manuscript. Only management data concerning growing seasons and fertilizer application rates besides the supply of sufficient nutrients in an additional scenario were harmonized (see Sections 1.3, 2.2, and 2.6), while models also of the same family still differed in their setups and parameterizations (Sections 2.1 and 2.2; Table S1-1). Following the suggestions of Reviewer 1 (comment 10), we are now treating also the EPIC-based GGCMs as separate models due to the selection of different subroutines. The harmonization hence serves for identifying remaining differences if key input data are eliminated as a source of uncertainty. As shown in the revised manuscript, this allows then indeed for identifying the impact of difference in biochemical routines, such as soil nutrient cycling and hydrology, especially if further harmonization is carried out as was done here for two of the models (Sections 2.7 and 3.5). The completely rewritten discussion (Section 4) now addresses differences in biogeochemical model parameterizations and their implications for yield estimates, model agreement and performance more specifically.

Please also note the supplement to this comment:
http://www.biogeosciences-discuss.net/bg-2016-527/bg-2016-527-AC1-supplement.zip

---

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

Printer-friendly version

Discussion paper