Geosci. Model Dev. Discuss., 7, C1954–C1965, 2014
www.geosci-model-dev-discuss.net/7/C1954/2014/

Geoscientific Model Development Discussions
Open Access

Interactive comment on "The Global Gridded Crop Model intercomparison: data and modeling protocols for Phase 1 (v1.0)" *by* J. Elliott et al.

J. Elliott et al.

jelliott@ci.uchicago.edu

Received and published: 9 October 2014

General comments: This paper presents a project's approach to global gridded simulations for the period 1948-2012. The paper should be a useful reference for both crop modellers involved in the project and more broadly also for other scientists that aim at using the project's public outputs for their analyses. The methods and data sources presented in the paper can also be of use to other researchers conducting regional or global-scale crop simulations. The paper provides a great deal of detail on many of the assumptions that will go into the project's simulations, including clear descriptions of weather and crop data. The GGCMI project is mainly an improvement over the work presented in the so-called 'fast track' (mainly Rosenzweig et al. 2014). My main concern is that the authors do not demonstrate the methodology, or even parts

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

of it. The paper is currently limited to showing the input data. This is fine, but maybe not enough for a scientific paper. For instance, one can think of some evaluation exercise of the Rosenzweig et al. (2014) model output over the historical period, using either the Iizumi et al. (2014) or the Ray et al. (2012) datasets, or both. This can provide an idea of whether there is scope for improvement in model skill through using better model inputs or scope for uncertainty reduction by 'harmonising' inputs. Taking advantage of the same simulations, authors can also show the type of extreme-event analysis that would be done. This can help the authors in framing / contextualising a bit better their objectives, and would improve substantially the paper. I suggest some revisions be made mainly targeted at removing ambiguities and better contextualising phase 1 within the project and the project's objectives more broadly in the context of climate change impacts research.

- The GGCMI is not necessarily an improvement but a follow up exercise to the fast-track which basically only reported on model differences. The objective of this paper presented here is not to describe the methodology of the analyses conducted in GGCMI and with GGCMI data, but to provide a clear description of the modeling protocol and the model input data provided by GGCMI. We describe data sets that will be used for model evaluation and show examples of how this evaluation could look like, but the intention of this paper is not to be a comprehensive methods section for the evaluation publication that is to follow in one or several following papers. Further, its generally not possible to use the fast-track outputs for the types of analyses considered here, because the "historical period" in the fast-track is just from climate model output rather than observation or even reanalysis-based weather. For this reason the results of the fast-track cannot be directly compared to observation-based yield estimates like Iizumi et al (2014) or Ray et al (2012). Indeed this is a significant motivation for the design of the GGCMI.

Specific comments: 1. Relevance / context of the project. GGCMI phase 1 will conduct global simulations of as many crops as possible for a historical period with four main

GMDD

7, C1954–C1965, 2014

objectives. Authors could expand a bit on the three-year GGCMI project so that the reader gets a clearer idea of how next phases will build upon phase 1. It would also be useful to see at least a brief discussion (in Sect. 6) of how this project overlaps / feedbacks from / contributes to regional assessments that are currently being carried out / funded by AgMIP itself or by other programs (e.g. CCAFS). Moreover, the context of these analyses (i.e. global gridded simulations) within the impacts research literature should also be stated (also see point 2 below).

- This is certainly a great suggestion. We have expanded the discussion on phases 2 and 3 of the project in order to provide greater context for this first project phase, and clarified to a greater extent how each phase will build off the ones before. We have also expanded discussions in section 6 to clarify how the outcomes of GGCMI are expected to facilitate other projects within and beyond AgMIP and ISI-MIP, including global and regional agro-economic and biophysical climate impact assessments.

2. Relevance / context of project objectives. It is not entirely clear, why are some of these four objectives being researched. While items (2) and (4) are clear overarching needs and/or knowledge gaps, the hypothesis and/or context behind item 1 should be stated more explicitly. More specifically, what new knowledge is expected to be generated by running models with harmonised and non-harmonised inputs? For item 3 (uncertainties) it is not clear which uncertainties or why do the authors choose to quantify these? is there evidence suggesting they may be a major source of uncertainty in yield hindcasts? On the input weather one can also think of bias correction of climate model meteorology? why are these not being researched (from a climate change perspective they may be at least as relevant)?

- The motivation for item 1 includes exploring how important varying assumptions on growing seasons and fertilizer inputs (or inclusion of nitrogen dynamics) actually are for simulated dynamics. Historic simulations allow for assessing how well observed variability can be reproduced by the models and how strongly this depends on assumptions on management. However item 1 also includes comparisons of some more fundamen-

GMDD
7, C1954–C1965, 2014

tal model choices, such as the method uses to calculate evapotranspiration within the models. In phase 1, we are performing a detailed intercomparison of different ET methodologies using the fact that some participating models (pDSSAT, pAPSIM, and the EPIC-based models) have the ability to simulate multiple ET methods with all other elements held fixed. This was mentioned only briefly in the initial submission, but that oversight has now been corrected and this example of deep model intercomparison has now been highlighted to clarify our motivations. Item 3: the uncertainties are to be derived from the differences between models and scenarios (weather datasets, management assumptions) also in order to facilitate a targeted attempt to improve model skills. The point is not to understand the uncertainty in yield hindcasts but to assess model skills from their ability to simulate historic yield dynamics and spatial patterns. The uncertainty in bias correction is certainly also an important one but we have put the focus on the different weather data sets available. We include, however, 2 raw reanalysis products that can shed some light on the general importance of bias correction, even though not on different methods of doing so.

3. L20-25 P4388: having in mind the four objectives stated at the beginning of Sect. 2 it does seem that running crop models where crops are not currently grown is unnecessary. Particularly for climate variability (obj. 4) and model evaluation (obj. 2) assessments. Maybe authors have a purpose for this (e.g. for further comparison to any future simulations that will be done in a follow up phase). However, as of now, why not just use some prescribed "crop mask" per crop and so in this way do not waste computational resources and facilitate further analyses? This is particularly important for northern hemisphere cereals such as wheat and barley whose climate requirements are unlikely to exist in large areas of the tropics. Vice versa for tropical crops not adapted to cold (e.g. cassava). The niches of the crops need to be maintained somehow. This brings confusion to the reader: for instance, in Fig. 4 (right) of this paper one can already see wheat in the Sahel.

- See also response to Beth above (point 2). We note that figure 4 (right) is produced

GMDD
7, C1954–C1965, 2014

not from simulation output but instead from national and sub-national observations compiled by Ray et al (2012). Additionally, the MIRCA dataset of crop covers that is used throughout the project does indicate that there is a small but nonzero amount of wheat grown in this region (see Figure 1 below).

Figure 1: MIRCA land-cover dataset for rainfed wheat area in sub-Saharan Africa. The global M3-crops dataset (Monfreda et al, 2008) shows a similar result).

4. L1-10 P4389: crop duration is a key output for understanding differences across models, particularly when these are driven by mean temperatures. All annual crop models should be capable of providing this as an output. In addition, perhaps authors should somehow indicate how many models (or by percentage) can provide each output.

- Indeed, crop modelers are asked to report planting and harvest dates, which allows for deriving crop duration (see Table 4). We can provide information on some models with respect to intended outputs, but those models that have merely indicated their interest have also not provided much information on what variables they will actually report. This information will clearly be reported in publications using the datasets provided by the crop models.

Technical corrections: 1. L5, P4386: unless described briefly (i.e. what it is and how is it different to GGCMI) a reference to AgGRID may confuse the readership.

- OK, will briefly expand the description of AgGRID (or possibly scratch it)

2. L21, P4386: consider using regional-scale process-based models. Hybrid may be too ambiguous.

- These models are not regional-scale models as they will be run at the global scale. Also, even though this may be true for some, not all “hybrid” models are developed for specific regions (e.g. Pegasus). The classification is certainly ambiguous, and its usefulness will have to be proven. Here we just want to highlight that we have field-scale

GMDD
7, C1954–C1965, 2014

models, land-surface/DGVM type models and other global gridded crop models, that we subsume under “hybrid” as they typically have a larger share of empirical relationships than field-scale or DGVM type models.

3. L22, P4386: ditto above, why not just use ’statistical models’, instead of ’purely empirical’? - Done, thanks.

4. L27, P4386: ‘modelling groups‘, rather than ‘modelers‘ - Correct.

5. L6, P4387: "such as" brings about some unnecessary ambiguity. Be specific. List clearly which uncertainty sources are being quantified. - Done, thanks.

6. L10 P4387: productivity, not production - Yes. Will be corrected.

7. L19-20 P4387: one would expect a relationship between the two measures (importance to food security / economies / livestock feed and number of models, or likelihood a model exists). It is likely that each criterion would yield the same list separately, hence it seems redundant to use both (with FS and/or economic importance being the independent variable). Besides, it seems reasonable to think that, as long as >=3 models simulate a particular crop (to allow for inter-comparison), the existence of many models should exert little impact on establishing the scientific problem / priorities. Also, the brackets on "(primarily global)" seem unnecessary.

- While there is certainly expected to be a correlation between the most modeled and most “important” crops, there are certainly circumstances where this is not the case. Many crops that are very important in economic terms (such as various cash crops, including coffee and tomatoes) or essential for nutrition in important regions (as e.g. sorghum, teff) are not modeled as frequently as some other crops.

8. Table 2: # models for priority 1 states 15-20 models. How can a crop achieve 20 individual model simulations when Table 1 lists 18 crop models?

- GGCMI is constantly growing and accepting new members and participants, so it’s somewhat difficult to say precisely how many models will contribute in any given phase.

GMDD
7, C1954–C1965, 2014

At least 2 new models have joined the group since initial submission with the intention of contributing results in time to participate in one or more paper for phase 1, so once these are added to the table the 15-20 estimate is more logical.

9. L18 P4388: "For the purposes of various analyses". Which analyses? if described in this paper please ref. the section. If not described in this paper then please do so, or state briefly what is meant by "various".

- OK, will do that. We have generally tried to make it clear that this modeling protocol lays the basis for many analyses, several of which have been scoped but not yet strictly planned in detail, for which we will try to provide suitable data.

10. L16 P4389: or maybe also to be able to interpret the differences in simulated yields?

- Certainly a good example of a future analysis not anticipated in advance would be the proposal and evaluation of a hypothesis of what is driving yield differences that has not yet been considered.

11 L18-20 P4391: This is unclear. While it makes sense to think of a growing season for comparability across models, observational datasets are generally based on the reporting standard of FAO, which uses whatever the countries report. In this scheme, yields reported in one year correspond to crops harvested in that year. It is not "artificial", as authors state. Authors are advised to cross-check their statement against the FAO reporting standard.

- According to the FAO glossary (http://faostat.fao.org/site/375/default.aspx), faostat yield estimates are usually produced by collecting production and area data and taking the ratio. For crop production, the definition in the glossary says the following: Crop production data refer to the actual harvested production from the field or orchard and gardens ... When the production data available refers to a production period falling into two successive calendar years and it is not possible to allocate the relative production

GMDD

7, C1954–C1965, 2014

to each of them, it is usual to refer production data to that year into which the bulk of the production falls. The procedure implemented by FAO is to assign the production to a given calendar year based on when that production is reported. In some countries this date can actually come significantly after the date of harvest. - Many crop models use a similar definition but this runs into problems when you're trying to compare among models or indeed when trying to compare to FAO. For example, in areas where harvest occurs near the new year, it may fall in some years in December and in other years in January. This often leads to calendar years with twice the normal production and other years with none. Clearly in this case assigning production strictly to the calendar year in which is falls is not the best option. Furthermore, models typically don't say much about when harvest of crops actually occurs, but instead only when the crops are matured. This is further complicated by the fact that FAO assigns production to a calendar year based not on harvest but instead on when the production data is reported to FAO, which as they note can "come significantly after the date of harvest". There is thus no consistent way to reproduce the FAO definition within a model protocol. The approach we have chosen comes as close as possible to being an unambiguous request to the model groups and leaves the difficult step of re-aligning outputs to match FAO to be done as part of the output processing pipeline, where different methods can be implemented and evaluated for relative performance.

12. It does seem a bit strange that the paper first describes simulation outputs and only after that describes the inputs.

- The goal of the paper is to describe output formats and protocols, not simulation outputs themselves.

13. L25-27 P4392: this statement is inconsistent with (actually contradicts) the purpose of the comparison of input meteorological datasets itself.

- Variable substitution is only required in very rare circumstances and for variables of secondary importance (long wave radiation may be the only example in fact, and its

GMDD
7, C1954–C1965, 2014

only used in a few models).

14. Table 11 should clarify whether ‘standard‘ (for wheat and barley) means spring.

- This has been clarified.

15. L6 P4394: sugarcane is harvested beyond 12 months in many places across the tropics

- Yes, but if we use the cropping calendar of MIRCA2000, sugarcane grows for exactly 365 days. For consistency and lack of better data with sufficient coverage, we stick to this. Also, cropping cycles >12 months would interfere with the annual character of agricultural systems that is embedded in many of the participating models

16. L13 P4394: LAI will not be zero for indeterminate crops

- True. For those we simply describe harvest dates and make no effort to adjust for maturity.

17. L3-12 P4394: it does seem like too many assumptions for areas in which no model evaluation can anyway be performed, and for which little scope exists for inter annual variability assessments.

- For various reasons described above, we want to produce a best guess for what the planting date and growing season length will be in each grid cell, even in grid-cells where a particular crop is not historically grown. We have tried to come up with a simple hierarchy for picking this best guess based on the data that is available at a given point.

18. L1-4 P4396: unclear whether this is done for each input meteorology dataset or using which met data?

- Yeah. That was criticized above as well. We should make clear that it should be done for one assuming that differences in temperature are not that severe to account for many days.

GMDD
7, C1954–C1965, 2014

19. L21-25 P4397: why has this been done? clearly, it will affect simulations of models that account for nutrient availability and/or uptake, mainly across the developing world. If this procedure is inconsistent with observations then what is the expectation with regards to model evaluation?

- See previous answers and also for the other review. This is for the extrapolation to currently uncultivated land and will thus not affect model evaluation. However for various purposes we need to produce a best guess for what management practices would be in a grid cell if a given crop were grown there.

20. Sect. 4.1. Perhaps it would be good to include some basic quality checking for the yield data (see for instance wheat in the Sahel, Fig. 4 right). In addition, FAOSTAT reported yields also have known issues.

- Yes that will be part of the evaluation study. Actually, strong disagreement with all models could be an indication of poor data quality in the reference data sets (although of course there a many other possible reasons for disagreement).

21. L17 P4399: "various analyses". Please specify

- Clarified.

22. Sect. 4.2.2. Detrending of FAOSTAT data may imply the need to detrend yield simulations as well, if climate change driven yield trends for the period analysed are observed in the simulations.

- Indeed, trends are removed from both the observation and simulation sets. For consistency, the same method is used to correct both (matching linear-detrended observations with linear-detrended simulations, etc.).

23. Sect. 4.2.3 be consistent with terminology: validation vs. evaluation. Validation suggests universality (not this case), hence it seems best to use the term evaluation.

- Agreed, thanks.

GMDD

7, C1954–C1965, 2014

24. L6-8 P4401: It is unclear how this will be achieved only with yield simulations and observations. You need an entire series of prognostic variables and measurements in order to conduct such an assessment. It also seems unlikely that regional-scale evaluation of yield simulations can drive model improvement. Far more detailed data are needed for such task.

- Agreed. But as a first step, we try to identify areas (crops, regions, events) where crop models performance is weak. Once these cases have been identified, we can try to find general patterns and supplement additional targeted analyses for these. The global gridded crop models are intended to work at regional scale, so an assessment should work at the scale of application and any model deficiency at the scale of application can certainly inform targeted model improvement.

25. L14 P4401: "stakeholder", please clarify / expand. - Clarified, thanks.

Interactive comment on Geosci. Model Dev. Discuss., 7, 4383, 2014.

GMDD
7, C1954–C1965, 2014

Interactive Comment

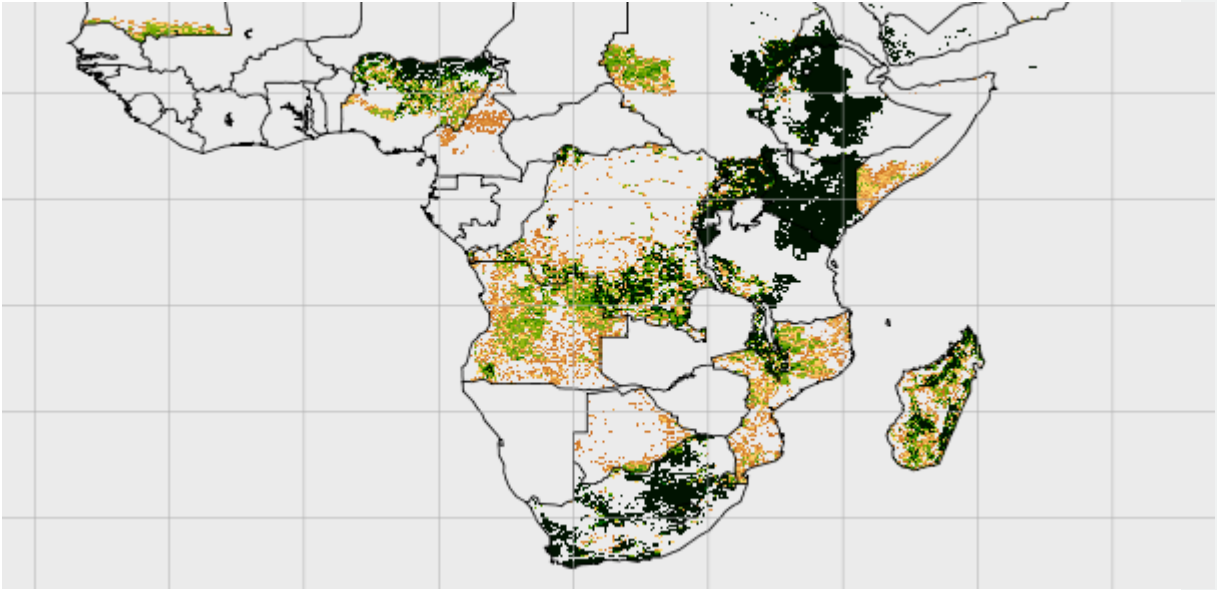

Fig. 1. MIRCA land-cover dataset for rainfed wheat area in sub-Saharan Africa. The global M3-crops dataset (Monfreda et al, 2008) shows a similar result).

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper