Response to Reviewer 2

R2C1: This paper attempts to improve the capacity of a global hydrological model H08 to simulate the yields of four major crops. The authors substantially improve the crop growth algorithm and add additional calibration procedures to aim at a better fit with national crop yields reported by FAO. The latter is arguably the main point of concern here. Calibration and validation were performed using two very similar reference datasets, both based on the long-term mean of FAO's yields. Due to this similarity, the validation shows great agreement with reference data. However, once the calibrated model is compared to long-term (not mean) annual yields, the agreement substantially drops and hence the model provides a weak representation of the actual historical yields. Another issue lies in the insufficient description of Materials and Methods. As a reader, it is difficult to understand the details of crop growth processes, simulation setup, and calibration procedure. Therefore, this section would greatly benefit from adding figures (e.g. crop sub-model scheme) and tables (e.g. main input data). Lastly, it would be interesting to see if the improved model can simulate the water cycle more accurately since H08 is primarily a hydrological model. Given mentioned-above, I would suggest major revisions.

Response: We thank the reviewer for the constructive comments. We have thoroughly revised unclear and incomplete points and incorporated all remarks. We address your general comments in view of three aspects:

1. Regarding the calibration

We believe it is successful in reproducing the long-term average crop yield (Figs. 1 and 2) and capturing the interannual yield variation (Fig. 3). For the majority of the top 20 countries, the anomaly magnitude of the calibrated simulations became closer to anomalies based on FAO data. We added text in the abstract (Lines 13–15) and the results section (Lines 261–264) to acknowledge the weakness of our current model in fully capturing the interannual variation of the historical yield:

Lines 13–15:

"*The capacity of our model to capture the interannual yield variability observed in FAO yield was limited, although the performance of our model was comparable with that of other mainstream global crop models.*"

Lines 261–264:

"Note that the calibrated model showed a similar performance to that of the default model in some countries (e.g., in USA, France, Ukraine, and Canada for maize) because the default simulations were already comparable to yield reported by the FAO, meaning that the calibration resulted in limited improvement (see Figs. 1a and 2a)."

2. Regarding the description of materials and methods

We added tables to specify the simulation settings (Table 1), adjustment of the harvest index (Table S1), and the main input datasets (Table S2).

We also added a reference to the schematic figure showing the biophysical process of the crop sub-model (Lines 108–109):

"A schematic figure that shows the basic biophysical processes of the crop sub-model is shown in Fig. 1b in Ai et al. (2020)."

3. Regarding the water cycle

There will be no change for the water cycle since we did not modify the hydrological modelling. Instead, as we repotted here, our enhancement can help to simulate the crop yield and food production more accurately. Therefore, our work is important to realize hydrology and crop growth coupled model to advance so-called food-energy-water nexus studies. Our model would be used to quantify the tradeoff between irrigation and crop production (e.g., Ai et al., 2021; Heck et al., 2020) and advance the analysis of water embedded in agricultural products and their international trade (e.g., Hanasaki et al. 2010; Dalin et al. 2017).

Specific comments

R2C2: Title: What does (crp.v1) mean? do you really need to include it in a title?

Response: As required by the journal, we need to specify the model version. This abbreviation is used to distinguish the model version.

R2C3: Line 9-14: The authors put too much emphasis on the "good consistency" (how do you define good?) with the long-term mean of FAO's yields. However, they fail to acknowledge the poor correlation when yields were analysed not as a mean but as a time series over the same period. The latter, one could argue, is much more important in the context of solving water-food nexus problems.

Response: Thank you. We agree and revised the text by further clarifying "good consistency" in terms of reproducing long-term averaged yield while also acknowledging limitations in terms of capturing the interannual yield variation (Lines 11–15):

"The simulated yield showed good consistency with FAO national yield. The mean biases of the major producer countries were considerably reduced to 2%, 2%, −2%, and −1% for maize, wheat, rice, and soybean, respectively. The capacity of our model to capture the interannual yield variability observed in FAO yield was limited, although the performance of our model was comparable with that of other mainstream global crop models."

R2C4: Line 20: Why do you refer to "food-water-land-energy nexus"? so far you only talk about water and food. Similar thing for line 345.

Response: Thank you. We changed "food-water-land-energy nexus" to "food-water". In our previous work, we worked on estimating energy crop for a truly "food-water-land-energy" nexus study (Ai et al. 2021), but we agree with you that we should focus on "food-water" in this study.

R2C5: Line 35: If you are trying to list all global crop models, then add the ones you miss from the recent round of ISIMIP simulations, e.g. ACEA, PROMET, SIMPLACE etc. (Jägermeyr et al., 2021).

Response: Thank you. We added the three modes, it now reads as follows (Lines 34–38):

"Many models have successfully incorporated the crop growth process and can simulate the global crop yield. These include LPJmL (Bondeau et al., 2007; Fader et al., 2010), GEPIC (Liu et al., 2007), PROMET (Mauser and Bach, 2009), PEGASUS (Deryng et al., 2011), CLM-Crop (Drewniak et al., 2013), PRYSBI2 (Sakurai et al., 2014), pAPSIM (Elliott et al., 2014), pDSSAT (Elliott et al., 2014), CROVER (Okada et al., 2015), ORCHIDEE-crop (Wu et al., 2016), PEPIC (Liu et al., 2016), MATCRO (Masutomi et al., 2016), SIMPLACE-LINTUL5 (Webber et al., 2016), and ACEA (Mialyk et al., 2022)."

R2C6: Lines 108-162: The list of formulas is helpful but it is difficult to follow. For example, the parameter "Ihun" is first used in Eq.1 but is only explained in Eq.9. I would strongly recommend starting Chapter 2.2 with a general description of the crop sub-model (preferably with supporting figure) and after that diving into details.

Response: Thank you. We restructured the section and described "*Ihun*" upon first mention (Line 119). We also divided section 2.2 into two sub-sections (2.2.1 Overview and 2.2.2 Basic algorithms). Section 2.2.1 provides a general description, and we also referred to the schematic figure of the biophysical process of the model (see response to R2C1). Section 2.2.2 provides the detailed equations.

R2C7: Line 165-174: you roughly explain the calibration of "blai" but you do not describe the calibration of "Harvest". As I understand, you first check if changing "blai" is sufficient to reach FAO's yield and only then you check if further changing of "Harvest" is needed, correct? In general, it is not clear how the calibration procedure works.

Response: Both *blai* and the harvest index were calibrated at the same time. Similar to the method of Fader et al. (2010), the harvest index was automatically adjusted according to the calibrated *blai*. Respective equations are now listed in Table S1.

R2C8: Line 243: Please indicate which models and time period you take from Jägermeyr et al. (2021).

Response: Thank you. We added this requested information (Lines 267–269):

"These countries were selected to make the data comparable with the latest global crop model intercomparison study by Jägermeyr et al. (2021), which includes 11 crop models for the period 1980-2010 (Fig. S10 in Jägermeyr et al., 2021)."

R2C9: Line 248-251: Jägermeyr et al. (2021) were not aiming at representing actual historical crop yields so it is not surprising that their data mostly demonstrates climatic signals.

Response: Thank you for the clarification. We understand the study more clearly now.

R2C10: Line 254: Double check that you actually take total precipitation and mean temperature for the duration of the growing season and not for the whole calendar year. Also, indicate how you weigh the grid cells here, is it according to harvested areas?

Response: We confirmed that we used the total precipitation and mean air temperature for the duration of the growing season. Furthermore, we used the harvest area for weighting.

R2C11: Line 276: Instead of referring to a "wide range of regions" you may refer to the percentage of grid cells with significant correlation. Same for areas without correlation.

Response: Thank you. We revised the text as follows (Lines 300–305):

"Using maize as an example (Fig. 8a), statistically significant correlations ($p < 0.1$) were observed in a wide range of regions (e.g., northeastern USA, southern Europe, northeastern China, southern Brazil, eastern Argentina, southern Africa, and eastern Australia) (Fig. 8a), corresponding to 31% of the total grid cells. Notably, there were also substantial differences in a considerable number of locations without statistically significant correlations ($p > 0.1$) (e.g., southeastern USA, western and central Asia, Brazil, and central Africa), corresponding to 69% of the total grid cells (Fig. 8a)."

R2C12: Line 294: There are ways to represent the historical trend in rainfed/irrigated harvested areas. To start, you can scale the total area to FAOSTAT and only then scale the yields. On top of this, you can include historical dynamics in overall rainfed and irrigated croplands as did for maize Mialyk et al. (2022).

Response: Thank you for providing this important study. We included it in the discussion (Lines 321–322):

"To overcome the problems associated with such an assumption, dynamic harvest area data at annual intervals, as generated by Mialyk et al. (2022) should be considered in future studies."

R2C13: Line 137, 343: Please avoid using the word "good" and use academic alternatives instead.

Response: We removed the word "good".

R2C14: Line 342: Where does the word "bioenergy" come from? You do not use it throughout the text.

Response: It refers to the enhancement of bioenergy crop simulation in an earlier study (Ai et al., 2021). However, we agree with the comment and removed text related to "bioenergy".

R2C15: Figure 2: It seems that for some countries the improved model shows the same numbers as the default one. Please explain why this is happening in Chapter 3.1.

Response: The calibrated model showed nearly the same interannual variation as the default model for some countries because the default simulations were already comparable to yield reported by the FAO, meaning that the calibration resulted in limited improvement. We further clarified this in the text (Lines 261–264):

“Note that the calibrated model showed a similar performance to that of the default model in some countries (e.g., in USA, France, Ukraine, and Canada for maize) because the default simulations were already comparable to yield reported by the FAO, meaning that the calibration resulted in limited improvement (see Figs. 1a and 2a).”

R2C16: Figure 8: Consider removing it. The message in lines 331-332 is very clear already.

Response: Thank you. We removed Fig. 8.

References:

Ai, Z., Hanasaki, N., Heck, V. Hasegawa, and T., Fujimori, S.: Simulating second-generation herbaceous bioenergy crop yield using the global hydrological model H08 (v.bio1). Geosci. Model Dev., 13, 6077–6092, https://doi.org/10.5194/gmd-13-6077-2020, 2020.

Ai, Z., Hanasaki, N., Heck, V. Hasegawa, and T., Fujimori, S.: Global bioenergy with carbon capture and storage potential is largely constrained by sustainable irrigation. Nat. Sustain. 4, 884–891 (2021). https://doi.org/10.1038/s41893-021-00740-4, 2021.

Hanasaki, N., Inuzuka, T., Kanae, S., and Oki, T.: An estimation of global virtual water flow and sources of water withdrawal for major crops and livestock products using a global hydrological model. Journal of Hydrology, 384(3-4), 232-244, https://doi:10.1016/j.jhydrol.2009.09.028, 2010.

Heck, V., Gerten, D., Lucht, W. & Popp, A. Biomass-based negative emissions difficult to reconcile with planetary boundaries. Nat. Clim. Change 8, 151–155 (2018).

Dalin, C., Konar, M., Hanasaki, N., Rinaldo, A., Rodriguez-Iturbe, I.: Evolution of the global virtual water trade network. Proc. Natl. Acad. Sci. USA., 109, 16, 5989-5994, https://doi:10.1073/pnas.1203176109, 2012.