The paper discusses some enhancements to the H08 global hydrologic model to simulating bioenergy yield over a history. The authors compare the results to previous assessments and some observed yield values around the globe. The paper is a good contribution, and I recommend its publication. However, the paper has several sections that require some additional clarity/details. Below I provide a detailed summary of some of these issues.

- It is not clear how this work builds on previous work by the authors (Yamagata et al. 2018 and Wu et al. 2019) or by the work of Trybula et al 2015.
- Better documentation of the methodology section to allow for reproducibility including the equations, and the explanation of the various parameters. A schematic would be also help.
- Sections 2.1 and 2.2 leave the reader wondering about the specifics of the two-step approach discussed, and how the adopted enhancements build on the previous approach. These two sections deserve more details of the methodology with greater levels of details that what is being offered. This will help the reader understand exactly how this work differs and builds on the two previous studies by the team, how to interpret the results and the difference between the 'original' and 'enhanced' versions of H08 (figure 3), how to interpret the various variables shown in Table 1, and to facilitate reproducibility.
- The paper shows some validation results for the rainfed module, and not for the irrigation module, but then show results for both when simulating both globally. The validation step for the irrigated module should be shown and discussed in the main text.

Other issues:

I would suggest shortening the title. How about something like "Simulating second-generation bioenergy crop yield using the global hydrologic model H08"

Line 4. Why is Miscanthus capitalized and italic but not switchgrass?

Line 7: 'enhanced H08' Doesn't H08 keep track of different version numbers that can be used here instead of calling something an enhanced model version?

Line 13: Add a sentence into the abstract to introduce the term BECCS if you are going to start the introduction section with this term. Preferably, I would suggest confining the framing around bioenergy crops rather than BECCS since the latter term never appears again in the text.

Lines 26, 30, 34: LPLmL should be LPJmL

Lines 30-32: It is not just LPJmL based on the following paragraph. It is also H08 based on the two recent publications using H08 (Yamagata et al. 2018 and Wu et al. 2019).

Line 34: change 'biogeny' to bioenergy

Line 41: Hanasaki et al 2008a/b are repeated twice in the list.

Line 44: the reference Wu et al. 2019 is missing in the list of references at the end.

Lines 49-50: I would suggest omitting the sentence "However, it is noted that the model performance for the simulated bioenergy crop yield was not validated at all" as an argument to justify the novelty of the work. I doubt the authors are claiming that the previous two studies using H08 with representation of bioenergy yield ignored properly validating the model and this study contributes this novelty. I would

suggest that authors replace this sentence with an explanation of how the new work builds on the two-step approach documented in the two previous papers (Yamagata et al. 2018 and Wu et al. 2019).

Line 61: 'The six sub-modules', You have not introduced what those six submodules are yet. I would start by listing them or at least list them in () right after this phrase.

Line 75: I would expand on this section to show the two-step approach here before talking about model enhancements in the next section (2.2). Even if those were presented in the two previous publications, I would at least include them in SI to make this manuscript a standalone piece.

Lines 76-85: I would suggest including all the equations and steps for how yield is simulated to shed more light about the method and to allow reproducibility of the approach.

Line 90: 'as an output item' Are you saying that can you simulate water consumption as a new output variable? It is not clear.

Line 91: 'Fifth, we fixed the bug in the original code'. What Bug? One could say 'we fixed a bug in the original code'. But this is so vague and does not really give the reader any additional information. I would suggest dropping the fifth point. Such details are best documented in SI.

Lines 105-110: can you mention the number of data points and years being used?

Line 115: what variable is being calibrated here? H08 simulates many output variables. How does the calibration process ensure that the adopted calibration process does not offer a gain in better matching one variable at the expense of another variable? For example, did the authors calibrate runoff first and then yield, or is it done all at once? If it is the latter, then showing some results on runoff would be necessary. I am not asking the authors to necessarily do additional work, but rather to better explain their approach.

Line 117: 'the enhanced h08'. Does this mean that the second simulation was only done for the enhanced model?

Lines 124-125: A bit unclear. Was the calibration done as a multi-objective optimization process to optimize both the RMSE and R values. For example, how do you decide an optimal parameter set when the two goodness-of-fit variables disagree? Figure 3 only shows RMSE, so I would suggest that you stick to this one and drop the R coefficient. Also, it is not clear if observed data is available for several individual years or only a single average year is available. If a time series exists, then I would suggest using goodness-of-fit measures such as Nash-Sutcliffe.

Line 137: 'because the few sites that were irrigated'. Please rephrase.

Line 139: 'previous reports' Please add citations to support this claim. The single sentence that comes afterward is insufficient. What about other parameters?

Lines 140-145: how does this work differ from Trybula et al 2015? This is not discussed in the intro. Also given that the adopted approach follows the SWAT implementation in Trybula et al 2015, and almost all of the parameters taken from the literature are also taken from Trybula et al 2015, would not it guarantee that you get similar parameter values for the other calibrated values to match those in Trybula et al 2015? What about other studies?

Lines 148-154: although the results are better than the original version, the results still seem to show a tendency to underestimate based on the results shown in figure 3.

Line 158: 'well at sites 1, 2, and 10' so how many sites are under irrigation? You should mention it here.

Lines 166-175: Did you drop the missing value from the significance test analysis (e.g., Finland in Fig 5d, Mongolia in Fig 5e)? I am still unsure whether the yield values from the other studies are average values over a particular period, and if it is the same period as in this study.

Lines 188-189: 'This can also be inferred from the validation results in Heck et al. (2016)' Please elaborate.

Lines 196-201: This information should appear earlier in the manuscript, so the reader is left wondering about such details. Also, if there is annual data from the other studies, then why not look at the timeseries instead of simply comparing the average value over a time period? To say a model can capture the long term mean over different basins is one level of validation, but to say that the model can also capture the interannual variability of yield from year to year, then this is a much more desirable level of validation.

Lines 203-220: This section comes as s surprise as it was not mentioned earlier as part of the framing of the paper in the front sections.

Line 206: It is confusing how many simulations were done in the study. The authors talk about two simulations twice, but are referring to different ones. I would suggest including an experimental design section as part of the methodology section to explain the different simulations to be conducted over a historical period (rainfed/irrigated, original/enhanced, …).

Lines 211-220: The validation results shown and discussed in the main body of the manuscript only talk about the rainfed simulations. It is unconvincing to skip the validation step for the irrigation module, and then show results and draw conclusions using that irrigation modeling capability. In this section, results from this study are shown, but they are not contrasted with estimates from previous studies.

Lines 223-233: Is this based on some aggregated regions, or on all the grids that belong to each climate zone? How do these results compare to other studies that were discussed in section 3.4? assuming this was based on a grid-level analysis, why not plot the results for all the grids and show a scatter plot (yield on the y-axis, and aridity or some other index that allows for distinguishing among the different climate zones on the x-axis)? This would allow the authors to fit a line to the data and talk about the results in a more compelling way.

Line 241: 'WUE, which is defined in this study as the ratio of yield to water consumption' This should have appeared the first time the term is mentioned in the manuscript.

Line 246: 'The WUE values for Miscanthus were higher than those for switchgrass, which is inconsistent with values in previous reports' Please add a sentence to articulate why?

Line 263: 'which was useful for optimizing bioenergy land with better consideration of water protection' – I am not sure what this means?

Lines 266: 'and our results are reproducible with the transparent parameter disclosed.' Just sharing the parameters sets does not guarantee reproducibility. I would suggest omitting that phrase.

Line 277: why was not this yield map used in the previous sections as part of the validation exercise? Also, I would suggest moving figure S7 out of SI and into the main text.

Figure 4: To be consistent with the black error bars, the blue/red ones should also reflect max/min. Also, why include all the years for observations? Should not these be for the years for which there is an associated observed yield value?

Figure 7: why is the y-axis for panel b flipped around as if the values should be negative? I would suggest keeping it consistent with the other two panels (0 at the bottom left corner, and the bar chart goes upward for positive values).

Table 1: Please add another column to define the different parameters and what they mean physically.