Response to Interactive discussion

Hydrology and Earth System Sciences (HESS)

Title: Assessment of Precipitation Error Propagation in Multi-Model Global Water Resources Reanalysis

Md Abul Ehsan Bhuiyan1, Efthymios. I. Nikolopoulos1 , Emmanouil. N. Anagnostou1 , Clement Albergel2 , Emanuel Dutra3 , Gabriel Fink4 , Alberto Martinez de la Torre5 , Simon Munier2 , and Jan Polcher6

We would like to thank the students for reading our submitted paper and providing insightful discussion and constructive suggestions. Below we provide a point-by-point response to these comments. Comments are in red and our responses in black font.

The main objective of the article is the assessment of both uncertainty in precipitation forcing and in the structure of several land-surface models by simulating hydrological variables. Methodology proceeds by studying the relative differences of three simulated hydrological variables by running five state-of–the-art models each forced by six precipitation datasets of various source. I think there is a strong need of capturing the relative influence of uncertainty from as well the datasets as the model structure to increase efficiency in hydrological predictions for water resources and to explore the possible usefulness of multi-model/multi-forcing ensembles. Therefore, the assessment of integrated structural errors would be an addition to the scientific literature on this topic and would definitely be a topic of interest for HESS audience. The paper is generally well-written in an understandable way, and the use of English is good. However, I my view the manuscript in its current state of is yet ready for publication and needs revision before it can be accepted. I hope the authors are willing to modify their manuscript, also taking into account the comments provided here. I feel the paper might benefit from a thorough restructuring to assure the manuscript meets the expected quality for HESS publications. Specific comments of three

Major recommendations

The introduction starts very clear and leads correctly to a certain problem statement. Although, the research question and aim of the paper which are following, are very broad and besides I do not get a clear view of the specific objective of the authors. The integration between precipitation error and model structure error is important and should be assessed, but the question raised in this paper needs a more precise aim. Momentarily, the aim of this paper leads to general results and a broad discussion and conclusion. For example, the conclusion partly states that the interaction between precipitation characteristics and different modelling schemes is very complex and the uncertainties in model simulations are due to precipitation and modelling structure errors. This conclusion is not a contribution to the scientific literature of this topic (Haddeland et al. 2011), while this subject has the competence to deliver a valuable product. To be more precise, in the introduction a problem is sketched which needs attention but the following query is too broad. In my opinion, the paper should focus on a smaller and more

specific subject to deliver an enhanced final product. The decision for coming to an alternative/specific aim is completely in hands of the original authors. Several suggestions to narrow this subject:

• Decreasing the number of precipitation forcings to one. As a consequence, more detailed results will be provided and a quantification of how the precipitation uncertainties behave or affect the different model-structures of the land-surface models. Via this way, the integration of the precipitation and model structure uncertainty will remain the main subject of the paper, but it will limit the scope.

• The other way around is also possible. By decreasing the number of land-surface models to one, the paper will show how the different precipitation forcings will influence the model structure uncertainty (in terms of hydrological variables).

• If, according to the authors, the essence of the article is damaged by reducing the number of forcings or models, the relative same procedure as before can be used. The disadvantage of this method is the huge amount of workload, because the results require a distinct end product, like quantifications of uncertainties, distinguished per model or per forcing

The second concern is the main methodology of simulation. Although the selection of the several diverse precipitation products and the five state-of-the-art models is proper, I do not understand why the one hydrological model is calibrated, but the four landsurface models are not. From my point of view, the outcomes of the simulations are only useful (in terms of considering model structure errors) when the models function in the same way, excluding the relative differences caused by alternative parameterization. This is also confirmed by Yin (2018), indicating that calibration of the models is necessary to use the same parameters while only the meteorological data is unique. To eliminate the effect of different parameterization, I expected the models to be calibrated to expose the uncertainties in model structures. Despite the fact that I am not considering myself an expert in the field of model calibration or validation, the main point here is the fact that this essential part of the methodology is rather unclear and additional information will be needed. Thus, if the authors are able to verify that the current parameterization does not affect the results of the exposure of model structure error, the explanation of this issue is very important to include in the paper. If the results are possibly affected and the land-surface models do contain incomparable parameters, the assessment of model structure uncertainty may be inaccurate and therefore cannot be identified as a certified result. Then I would recommend to perform model calibration and consequently present the improved results. For inspiration of model calibration, I found very intriguing papers, such as the paper of Beven and Binley (1992) about the GLUE framework, and the paper of Clark (2008) about the modular framework FUSE, focused on models tructure. The last issue is the presentation of the results. Although the approach of the results is very clear, the display of the graphs is quite cluttered and the amount is way too much. This issue relates to the first point of the review, providing the paper a more distinct aim. If the results would be presented more specific, the exact objective will be targeted

much more efficient. At this moment, one does not know where to look for between 16 different graphs on one page and it gives the impression the reader needs to search for the results himself.

I would recommend to show only results that are specifically relevant to answer the research question, to prevent clustering of graphs and figures. Just use in the result section no more than one or two graphs of each sort of visual representation to support the corresponding findings. This recommendation does not even revolve around

the possible change of the research question; even if the research question does not change, the authors should think of other ways of presenting their results.

References

Beven, K. Binley, A. (1992). "The Future of Distributed Models: Model Calibration and

uncertainty Prediction". Hydrological Processes, 6, 279-298.

Clark, M. (2008), "Framework for Understanding Structural Errors (FUSE): A modular

framework to diagnose differences between hydrological models". Water Resources

Research, 44(12).

Haddeland, I. (2011). "Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results". Journal of Hydrometeorology, 12(5), 869-884.

Yin, Z. et al. (2018). "Comparing the Hydrological Responses of Conceptual and Process-Based Models with Varying Rain Gauge Density and Distribution". Sustainability,10(9).

Few studies have been dedicated on the analysis of the integrated impact of both forcing and model uncertainty on hydrologic simulations and from the existing ones most of them were focused on a single hydrologic variable such as streamflow/evapotranspiration. So, this paper uses the multi-forcing/multi-model experiment to address the following research questions:

1. How does the precipitation uncertainty propagate through the multi-model hydrologic simulations?

2. What is the relative importance of precipitation vs. modeling uncertainty on the simulation of key water cycle variables (surface/subsurface runoff and ET)?

3. What is the spread of the precipitation uncertainty in simulation of hydrological variables and how this depends on model type?

As mentioned above, this paper presents a unique precipitation-to-hydrologic simulations error analysis based on different hydrologic variables, multiple models and multiple precipitation

datasets, to evaluate the role of uncertainty in precipitation forcing relative to modeling error. For this purpose, we considered multiple precipitation datasets and a number of global and land surface hydrologic models, which led to a very comprehensive error propagation investigation. At the same time in our revised version we will further expand the discussion on model description, which will clarify the rationale of using both calibrated and uncalibrated hydrologic models.

General comments

One of the keywords of this paper is "model structure" and its respective uncertainty. However, the definition of this keyword is not thoroughly described and therefore different interpretations are allowed. In addition, there is a fine line between the concept of model structure and their corresponding parametrization. Thus, I would recommend to sharply define this keyword and so delineate its meaning for this paper. I would recommend to add a line to the section concerning the study area why this area has been chosen. This is especially needed because in the discussion section (line 1-4) the authors insinuate that one of the hydrologic variables (evapotranspiration) is not the best measure of sensitivity for this study area. This is quite logical, because the study area is semi-arid and the amount of evapotranspiration is water-limited instead of energy-limited. Although the location of the study area is not damaging the research, in my opinion it would help to support the choice in this case.

Thank you for your suggestion. In the revised paper we will clarify these issues.

Further small remarks

Page 4, line 20: Change "hydrologic variable" to "hydrologic variables".

Thank you. It will be corrected in the revised paper.

Page 5, line 15: Change "semiarid" to "semi-arid" (for consistency, because at page

18, line 2, also "semi-arid" is used)

Thank you. It will be corrected in the revised paper.

Page 6, line 4: remove the capital letter of 'Land surface model"

Thank you. It will be corrected in the revised paper.

Page 7, line 11: Change "3hourly" to "3-hourly".

Thank you. It will be corrected in the revised paper.

Page 8, line 15: Add comma after "column".

Thank you. It will be corrected in the revised paper.

Page 12, line 13: Change “NCRMSE (variables)” to “NCRMSE (hydrologic variables)” or “NCRMSE (simulated variables)”.

Thank you. It will be modified in the revised paper.

Page 17, line 23: Change “result” to “results”.

Thank you. It will be corrected in the revised paper.

Page 18, line 18: Change “ORCHIDDE” to “ORCHIDEE”.

Thank you. It will be corrected in the revised paper.

Page 32: Please add the full description for the graph, as on page 31.

Thank you. It will be explained in the revised paper.

Page 33: Please add the full description for the graph, as on page 31.

Thank you. It will be explained in the revised paper.

Page 37: Description of the Taylor diagram of SURFEX: change “maen sqaure” to “mean square”

Thank you. It will be corrected in the revised paper.