Answer to Comments to the Authors by Reviewers of "HESS Opinions: Participatory Digital Earth Twin Hydrology systems (DARTHs) for everyone: a blueprint for hydrologists' by R. Rigon et al.

Comments by Dr. Uwe Ehret

C 1- Dear Editor, dear Authors, I have reviewed the aforementioned work. My conclusions and comments are as follows:

1. Scope

The article is well within the scope of HESS.

2. Summary

In their opinion article, the authors describe both the necessity and the key aspects of moving hydrological modeling from their current state towards Participatory Digital Earth Twin Hydrology Systems (DARTHs). Key aspects of such systems include i) flexible coupling to many sources of data available in standardized formats , ii) platform- and language-independent implementation, iii) modular design facilitating recombination of model components. Each such aspect is explained in detail in a separate section and briefly summarized in the conclusions.

3. Evaluation

Overall, this article is a relevant and timely contribution to the discussion about the future of earth system modeling in general and hydrological modeling in particular. It discusses many current hindrances towards efficient, integrated and collaborative modeling and names key elements of a framework to overcome these hindrances. So I welcome the article for its content,

A -1 - We thank the reviewer for the appreciation of our paper and we will do our best to accommodate his valuable suggestions.

C 2 but I do have several suggestions to improve its presentation:

Key messages: For a reader not strongly familiar with the current state of development of Digital Earth (DE) and Digital Twin Earth Models (DT), only after reading the entire manuscript it becomes clear what the differences to current hydrological modeling practice are, and consequently what the authors aim at with their article. A reader will be able to follow the arguments in the paper much better if this is made clear from the beginning. I therefore suggest the following changes to the manuscript:

- Include in the introduction a clear definition (purpose, key structural and organizational elements) of DE's, DT's, DARTH's and current hydrological models. Make clear where the latter fall short of qualifying as DARTHs, why it is necessary to elevate them towards becoming DARTHs, and what it takes to do so. A comparative list, or an explanatory chart can support this (maybe directly pointing to the sections

in which the individual topics will be discussed in detail). From such an introduction which provides the overall picture, it will be easier for the reader to connect to the individual sections.

A 2 - We will try our best to consider the Reviewer's suggestions in revising the introduction and to provide a better overview in the Abstract. However, we must also be careful not to overwhelm the reader at the outset with an overload of information in the Introduction. . Certainly, we will provide the required definitions and a short overview of the subsequent sections in the Introduction, as well as a Glossary explaining the main technical terms appearing in the paper.

C3 - At the beginning of each topical section (sections 2-9), include a very short overview of the section content.

A 3- <u>We will probably do this at the end of the Introduction</u>

C 4 - Abstract: From reading the abstract only, I did not have an idea about what to expect from the article, and what its key messages are. For example, I did not know how to interpret the term "model as commodity", which is not standard terminology in hydrology. Please rewrite, taking into consideration my previous comment about key messages.

A 4 - Yes, of course. <u>We will modify the Abstract to give a clearer overview of the paper</u>.

C 5 *- L271: I partly agree with the statement "The peer review process is ineffective at sniffing out poor model validation", but while it may be far from perfect, I'd say it is still the best we have. Please explain how in a DE/DT/DARTH world quality standards are maintained better than by peer control.*

*A 5 - The statement was certainly not aimed against the peer review process itself. We recognize, however, that even when code is Open Source it is difficult to find mistakes in models. In fact, we would go so far as to say that it is impossible without: a) proper software organization; b) simulations setup tracking and recording; c) clean code writing; and d) sufficient further documentation. Open Source code, however, can be inspected and errors potentially caught by third parties. The above characteristics in principle are not required in a DE/DT but we require them in DARTHs to improve the science check and reproducibility. <u>We will modify the text by adding "alone":</u> "The peer review process **alone** is ineffective at sniffing out poor model validation"*

C 6 Section 7: The section header just mentions "reliability", but in the text more aspects of model performance are discussed. I suggest widening the section header, and discussing robustness, reliability, reproducibility and realism (as mentioned in L260).

A 6- We think that enlarging the scope of the section beyond the aspects already analyzed would require another paper. Reproducibility is discussed throughout the paper and realism

of the model is certainly related to what was discussed in the section . We might consider changing the title of the section. <u>However, we will give a clearer definition of "robustness, reliability, reproducibility and realism" in the revised text and we will indicate where we have discussed the topics.</u>

C 7 - Section 10: The conclusions are written in a rather hasty and bullet-point style, which devaluates the otherwise interesting and comprehensive paper. Please spend some more effort in writing a coherent and standalone summary and conclusion.

<u>*A 7 - Conclusions will be revised in the next version of the work.*</u>

C 8 - Appendix A: Reading Appendix A, it was not clear to me why this section was moved to the appendix rather than being another section in the main text. I'd say the topic of proper coding is just as relevant for developing DARTHs as the other topics in sections 2-9. I suggest moving it to the main text.

A 8 - It would be easy to move this Appendix back to the main text. However, its character, in our opinion, is a little different from the other sections, and we prefer to keep it where it is.

Minor points

- L72: What is meant by "reasonable colour maps"?

Here we meant maps that show plausible patterns, obtained by using models that are solid and reliable from the scientific perspective, broadly used and tested, and that present few or no problems from the end-user's point of view. This sentence will be rephrased as " plausible patterns".

- L82-84: It is unclear at this point what is meant by "tight, black-box models", or rather how they differ from the previously mentioned models. Please clarify.

Black-box models are those models in which the internals are hidden in a "black-box", and the only parts that can be viewed are the inputs and outputs.

- L89-90: The meaning of the sentence ("At present, …") is not clear to me. What is meant by "modeling panorama" here? Please clarify.

In this sentence we mean that there are a lot of models available with different levels of complexity, making it difficult to simplify the matter. We changed "modeling panorama" into "models variety"

- L203: For me, the user's names are unfortunately not self-explanatory. Please add a short explanation similar to the explanation of roles.

Users are actually the runners, players, viewers and providers. The sentence is misleading and it will be dropped.

- L353: Please explain the "information hiding principle", because it seems to contradict the general paradigm of openness and accessibility advocated throughout the paper.

Here "information hiding principle" refers to the encapsulation principle, which in object-oriented programming prevents direct access to objects by clients to avoid exposing hidden implementation details or violate state invariance, maintained by the methods. This principle is not preventing the openness and accessibility, since it is a runtime process. However, we will add a glossary to define all these concepts.

- Caption of Fig. 5: What is meant by "IT"? Also, please add the channel network to the upper map for better connection to the lower map.

"IT" is a typo: it should read "It". The network will be added to the upper plot.

- *Table B1 is very helpful. Please mention it already in the introduction, such that the reader can make use of it while reading the paper.*

It will be done in the next version of the work

- Overall, there are many typos in the manuscript. Please check and remove.

A careful editing will be done on the work.

C 9 -I have another question to the authors related to current hydrological modeling and hydrological modeling in a DARTH environment. There is no need to address this in the manuscript, but I am interested in the authors's opinion. The question is related to emergence as larger-scale phenomena arising from strong interactions on smaller scales. Current models build for a specific purpose and a specific scale/resolution often either make implicit use of emergence (by taking emergence as a given, directly representing the emerged phenomenon) or, because it may be relevant, setting the model up in a way that emergence may actually happen as the model runs (e.g. emergence of convective thunderstorm cells in convection-permitting atmospheric models). The latter often requires a well-attuned choice of time-stepping, spatial resolution, numerical scheme, variable precision, processes represented, etc. With a free and modular combination of model components in a DARTH environment, we may miss such well-attuned combinations, which may lead to not only

incrementally but fundamentally different model behavior. How can this be taken care of in DARTH environments?

A 9 - The question raised by Dr. Ehret is important. In DARTHs constructions there are two aspects to be aware of. One is the infrastructure: the rules and the general principle of the organization of the infrastructure should be such that the infrastructure is agnostic to any science content. The other aspect is the science content.

In being agnostic, the DARTHs infrastructure is designed to accommodate as much as possible any science solution and allow its testing. The idea is to contrast the "hammer-nail" attitude (If the only tool you have is a hammer, you tend to see every problem as a nail.), present also in research, and to provide a very adaptable infrastructure in which any problem can find its proper solution (and not a solution that is shoehorned to the problem). The flexibility of infrastructure serves to test alternative hypotheses fairly and compare them against data evidence. The emphasis on tools establishing degrees of uncertainty is functional to selecting the more successful hypotheses, whichever they are, and eventually exclude those proving to be less reasonable.
Therefore DARTHs can evaluate any hypothesis, including the ones that involve changes in scales of the analysis.
Going to the science questions, we recognize that a statistical mechanics of hydrological phenomena is largely missing; however, the matter of emerging properties in hydrological models is the object of investigations that DARTHs could and should allow more easily than with traditional modeling. There is no problem in performing all the analyses that a sharp researcher would do with traditional tools. The Developers can implement new parameterizations more easily, the Linkers can explore the pool of model components, and the Users investigate the parameter space. The presence and use of tools to assess the error of estimation should eventually raise red flags when something is not properly modeled and warn the researcher to rethink the whole modeling process at their appearance.
One aspect that we probably did not stress enough is the fact that DARTHs impose a certain systematicity in the analyses (DE were said to be an "organizing metaphor"), the current absence of which is one of the obstacles to assessing laws of general validity. Recognizing emerging patterns and features should be easier when dealing with a multiplicity of catchments and from the many points of view that DARTHs promote.
DARTHs obviously do not substitute human creativity in formulating hypotheses, they only offer a wider range of tools to implement and test them appropriately.

Yours sincerely,

Uwe Ehret