

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
<https://doi.org/10.5194/hess-2021-644-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on hess-2021-644

Markus Hrachowitz (Referee)

Referee comment on "HESS Opinions: Participatory Digital eARth Twin Hydrology systems (DARTHs) for everyone – a blueprint for hydrologists" by Riccardo Rigon et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-644-RC1>, 2022

I have really enjoyed reading this manuscript, although I do not necessarily agree with all of the points made. As a very welcome but (unfortunately) rare case, this Opinion piece develops a broad visionary perspective of what could be a very valuable step for the development of scientific hydrology as core part of the Earth System Sciences. The authors do not only formulate their vision as a mere "wish list" but they also attempt to provide an outline of necessary major steps to be considered and potential challenges to be met along the way.

I have a few observations and suggestions the authors may want to consider, as I think they may be helpful to strengthen the impact of their work.

- This manuscript has been submitted to a hydrology journal. I therefore assume that the target audience envisaged by the authors are hydrologists and scientists/engineers from related fields.

As such, I suspect that many of our colleagues including myself may not be in detail familiar with some of the very technical and detailed computer science jargon used in the manuscript (some of them will be, of course) – in particular in sections 5 and 6. This may potentially also limit the level of appreciation and impact of this work. This would be unfortunate, really.

I think there are two alternative ways of dealing with that issue. Either, the authors

rework these heavy-jargon parts to make their language more accessible to a wider audience. Or the authors invest a bit more effort in more detailed explanation of the jargon terms, to allow the average reader to better follow their argument.

- The authors cover the aspect of technical steps and challenges in a very exhaustive way. This is very welcome and necessary. However, I also believe that the vision for the development of DARTHS can be further strengthened by outlining some of the questions, steps and challenges that will arise from an organizational perspective. This could include questions such as: which type of organization is necessary for the development, hosting and maintenance of such a system? How can a decision process in the (further) development of DARTHS be designed? Who decides what? Can/should DARTHS be non-commercial or does it need to be designed as a commercial endeavour? Or in other words: who can afford it, how can financing look like? Who is responsible for quality control of items added by users? Who is responsible for avoiding misuse and misinterpretation of the models/data by non-specialists (e.g. players, viewers and to some extent perhaps also runners)? See also some aspects in Weiler and Beven (2015) from the perspective of a community model.

- To grasp the context of DARTHS and its evident differences to other previous and ongoing initiatives that are currently state-of-the-art in our discipline (e.g. modular modelling frameworks), it will be helpful for the reader if authors provided a bit more detail in the discussion of similarities/differences with at least a few very **specific** other modelling frameworks. This could include comparisons with modular modelling frameworks at various levels of complexity such as SUPERFLEX (e.g. Fenicia et al., 2011) or SUMMA (e.g. Clark et al., 2015) and further extend to more versatile tools such as the very recently introduced eWaterCycle platform (Hut et al., 2021) – which is, in my understanding already quite a large step towards DARTHS. I believe a very simple table in which it is indicated which of the currently available tools already tick which boxes and which additional boxes DARTHS could tick.

- The language becomes a bit too informal in parts of the manuscript and could benefit

from being more precise to avoid ambiguities.

Detailed comments:

p.2, l.30: I suggest to replace "understanding" with "describing", as "understanding" is part of the discovery process and thereby meta-science. Whether or not you personally understand something is not really relevant (and there is of course no "collective" understanding). At one point something clicks in your brain but how is that relevant for other people? In other words, it remains something very subjective (and thus the opposite of what science should be).

p.2, l.39: Should probably read "Space Agencies" instead of "Spatial Agencies"

p.2, l.42ff: the use of "top-down" and "bottom-up" may generate confusion as they are typically used for very specific modelling strategies in hydrology/environmental sciences. Perhaps helpful to use a slightly different terminology here.

p.2, l.43ff: I found this a statement that is a bit too sweeping, generalizing and pessimistic. There are many research groups that actively work on model development/improvement. And any other research group that does not, is of course free to start working on this anytime. It reads as if these poor people are forced to use models imposed onto them by some higher force. Perhaps good to tone it down a bit.

p.3, l.61: do you really intend to already "answer" these questions? This seems a bit too

ambitious and restrictive. I think can be reformulated into something like "outline potential ways forward"

p.3, l.65: the term "certified" seems a bit awkward here. Not clear what you mean to say here.

p.3, l.72: what is meant by "reasonable colour maps"? Beautiful maps or maps that show plausible pattern? Similarly, what is meant by "...have no...issues"? how do you define issues? Depending on the definition, one could equally say, that *all* models have a lot of issues. Please rephrase.

p.3, l.74: I strongly disagree. That is what for example the many recent modular frameworks are aimed at.

p.3, l.79ff: Meaningful classification of models is indeed tricky but I believe the taxonomy provided here does not really capture the main differences in model features. The main differences between models, as we argued in Hrachowitz and Clark (2017), are the level to which physical constraints are imposed. For example, typically data-driven/statistical/machine-learning models (notwithstanding some recent developments) have not even imposed conservation of mass. Conceptual/reservoir -type models at least satisfy that constraint and work with a few additional process assumptions. The level of process representation then increases towards models, such as ParFlow which of course have much more detailed process representations. Therefore I would rather refer to all models that use at least some process assumptions as process-based on a gradual spectrum. In addition, I believe referring to lumped model implementations here can also spark some confusion. No matter which model type is used – it can be implemented at any spatial resolution. If this is justifiable is of course a completely different question. Perhaps try to reformulate this paragraph.

p.4, l.86: not clear what is meant here.

p.4, l.89: what is meant by "panorama"?

p.4, l.90: System complexity emerges to quite some degree from variability and heterogeneity. They are therefore intimately linked. However, the way it is expressed here gives the impression that "complex" and "complicated" constitute some sort of dichotomy, as in "on the one hand and on the other hand", while it should rather be that one follows from the other.

p.4, l.100: I do not really understand that statement. Of course ML can be "investigated". Why should this not be possible? It is a human-made construct. As such it can be adjusted but also looked into. I guess you mean until recently it was difficult for non-experts to analyse what is happening in the code of ML models. Please make this clearer. In addition, I do not believe that ML can have "knowledge". Makes it sound like a conscious entity, which it is (to my knowledge) not yet. Please rephrase.

p.5, l.134: this is not unique for the US. Environmental data from many European countries are also publicly and readily available. For example, Austria (e.g. <https://ehyd.gv.at/>), France (e.g. <https://www.hydro.eaufrance.fr/>), Germany (e.g. www.dwd.de), UK (e.g. <https://archive.ceda.ac.uk/>), and many others.

p.6, l.166: not clear what is meant by "binding".

p.6, l.178 (and elsewhere): please clarify the meaning of "seamless" here

p.8, l.212ff: I found this paragraph very difficult to follow and I am not sure what the authors try to express here. Perhaps helpful to reduce jargon or to explain in a bit more detail.

p.9, l.261: perhaps replace "reality" with "real world observations". In addition, please specify what is meant by "internals".

p.9, l.262: but this needs to be a very detailed knowledge of the simulation set-ups as demonstrated by Ceola et al. (2015) and generally argued to be impossible by others (e.g. Hutton et al., 2016). Please tone down and reformulate.

p.10, l.274: what is meant by "building tools"?

p.10, l.275: what does "...to certify the providers of models..." mean and entail?

p.10, l.283: please specify "all they need"

p.10, l.285: what does the "prepared simulation" include? Calibration set-up? Results? In addition, what is meant by "governed"?

p.10 or elsewhere: I am not sure where this fits in, but one aspect that seems crucial to me is the definition of the smallest, unchangeable building block of models in the entire system. What could these be? Can there be multiple? Who decides on that? Can users (e.g. Runners) just add such building blocks and/or specific parameterizations (as in reality we have no idea which parameterizations – i.e. equations, not parameter values! – are most suitable where/under which condition/at which scale/etc. see e.g. recent analysis by Gharari et al., 2021)

p.11, section 5: although I have a fair share of model development/coding experience, I struggled with the entire section. Frankly, I could not follow it. In particular, it was difficult to grasp what the subtle (or perhaps for the specialist not so subtle) differences between the five classes MaaA, MaaT, MaaS, MaaR and MaaC are and what follows from these differences. For example it would be very instructive and helpful if you could let the reader know into which class different existing models, modular frameworks and platforms fall (e.g. SUPERFLEX, HYPE, SWAT, SUMMA, eWaterCycle)

p.12, l.326: The role of the provider remains quite vague. Is this the data provider? Is this the developer of the model concept/idea? Is it the developer of a model code that is based on a specific model concept/idea? Is this somebody completely different? Also the term "policies" is unclear here.

p.14, l.365: "invasiveness"??

p.14, l.380: "...not confined to convey science from a single discipline." Sounds awkward. Please rephrase

p.14, l.384: what are "fake" models? Models that do not exist? Please rephrase.

p.14, l.388: "under the hood". Please rephrase.

p.15, l.462: "some conditions" is quite an understatement. With our currently available observation technology *most* process dynamics and system properties (e.g. soil hydraulic conductivities) are unknown at most locations during most of the time – in reality we have no idea of the spatial covariance fields of most of these quantities. Instead and to deal with this problem we make sweeping assumptions about this missing information and thereby we very likely upscale homogeneity instead of heterogeneity.

p.15, l.464: well, not only data errors, also model structural errors can and do result in parameters that do not reflect real world system properties.

p.19, l.480: is the range of results really that restricted? How is it then possible that

different models exhibit considerably different (internal) behaviours (e.g. Bouaziz et al., 2021)?

p.19, l.480ff: "some type of warning": this is extremely relevant and deserves some more consideration and detail in the text.

p.19, l.483ff: perhaps also good to refer to the exchange between Nearing et al. (2016) and Beven (2016), which is very reflective of these issues that are yet unsettled.

p.21, l.529-548: very interesting and important ambition!!

p.22, l.555: as recently also demonstrated by e.g. Gharari et al. (2021): given the limited observations we have relative to the size and complexity of our systems, process-based (i.e. "conceptual" and "physically-based") models can too restrictive with their assumptions on the type/shape of functional relationships.

Above I have added quite a few references of our group. Please see them as mere examples and suggestions. It was only done for convenience (easier for me to find our references than those of other groups). Needless to say that many other groups work on similar topics and their references may be more suitable. Therefore, please do not feel obliged to use the references suggested here.

Best regards,

Markus Hrachowitz

References:

Beven, K. (2016). Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, 61(9), 1652-1665.

Bouaziz, L. J., Fenicia, F., Thirel, G., de Boer-Euser, T., Buitink, J., Brauer, C. C., (2021). Behind the scenes of streamflow model performance. *Hydrology and Earth System Sciences*, 25(2), 1069-1095.

Ceola, S., Arheimer, B., Baratti, E., Blöschl, G., Capell, R., Castellarin, A., ... & Wagener, T. (2015). Virtual laboratories: new opportunities for collaborative water science. *Hydrology and Earth System Sciences*, 19(4), 2101-2117.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., ... & Rasmussen, R. M. (2015). A unified approach for process-based hydrologic modeling:

1. Modeling concept. *Water Resources Research*, 51(4), 2498-2514.

Fenicia, F., Kavetski, D., & Savenije, H. H. (2011). Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development. *Water Resources Research*, 47(11).

Gharari, S., Gupta, H. V., Clark, M. P., ... & Savenije, H. H. (2021). Understanding the information content in the hierarchy of model development decisions: Learning from data. *Water Resources Research*, 57(6), e2020WR027948.

Hrachowitz, M., & Clark, M. P. (2017). HESS Opinions: The complementary merits of competing modelling philosophies in hydrology. *Hydrology and Earth System Sciences*, 21(8), 3953-3973.

Hut, R., Drost, N., van de Giesen, N., van Werkhoven, B., Abdollahi, B., Aerts, J., ... & Weel, B. (2021). The eWaterCycle platform for Open and FAIR Hydrological collaboration. *Geoscientific Model Development Discussions*, 1-31.

Hutton, C., Wagener, T., Freer, J., Han, D., Duffy, C., & Arheimer, B. (2016). Most computational hydrology is not reproducible, so is it really science?. *Water Resources Research*, 52(10), 7548-7555.

Nearing, G. S., Tian, Y., Gupta, H. V., Clark, M. P., Harrison, K. W., & Weijs, S. V. (2016). A philosophical basis for hydrological uncertainty. *Hydrological Sciences Journal*, 61(9), 1666-1678.

Weiler, M., & Beven, K. (2015). Do we need a community hydrological model?. *Water Resources Research*, 51(9), 7777-7784.