

Interactive comment on “HESS Opinions: Repeatable research: what hydrologists can learn from the Duke cancer research scandal” by Michael N. Fienen and Mark Bakker

Michael N. Fienen and Mark Bakker

mnfienen@usgs.gov

Received and published: 13 June 2016

We thank Dr. Bellin for contributing to the discussion about our opinion piece. We are pleased that Dr. Bellin generally found our comments of interest and we welcome the suggestions for potential improvement of our presentation.

Considering uncertainty We are definitely advocates of considering uncertainty in all scientific endeavors. However, we have the opinion that discussion of uncertainty is really a topic in itself and not closely enough related to our main topic to add extensive discussion of it in the opinion piece. However, we will make strides to clarify (also in response to other comments) that repeatability and reproducibility

C1

are different. What Dr. Bellin identifies as a shortcoming by not accounting for the role of parameter and epistemic uncertainty really points to the difference between the two. Being able to repeat analysis and modeling with a specific parameter set may seem trivial since it doesn't consider the uncertainty of the process, but it is a necessary and often difficult to accomplish step! Even stochastic approaches should be repeatable and ensembles of parameter sets and resulting forecasts can be carried forward using scripts. This does not guarantee reproducibility in the case of epistemic uncertainty as it influences subjective decisions about the data/models, as another group with another model may come to different conclusions. If this is the case, it is crucial that the published results can be repeated, as it can at least be concluded that the difference is not due to errors in the published results but (likely) due to epistemic uncertainty. This highlights again that reproducibility is still an important issue that is not given enough attention in the hydrological sciences. We will clarify this in the revised version of our paper and highlight the importance of repeatability in the case of epistemic uncertainty.

The more general issue is to include more details than "letters" style articles Indeed, there are multiple aspects to how more detailed information leading to greater repeatability can be incorporated into scientific discourse. This was also an issue raised by Dr. Cirpka. However, we chose to highlight the response of the medical research community to the Duke Cancer Research scandal, being to require not just detail in writing, but an executable path through analysis via scripting. We use this as an example rather than insisting on this as the only solution. We are glad that the result has been a vigorous discussion so far and we will incorporate more about the general issue of repeatability in the revised manuscript.

Scarcity of hydrologic data It is true that the example from omics often are cases with large datasets that must be trimmed while in hydrology data are often scarce so trimming is less an issue. We can clarify this in our paper. However, the anal-

C2

ogy between the fields is more basic in our view. Whether the issue is trimming a large omics dataset or interpreting noisy and sparse hydrologic data, in both cases subjective decisions must be made about suitability of data. Since they are subjective, other researchers must be able to understand, assess, and, possibly, overrule such interpretations. By clearly documenting them in a scripted path through the analysis, other researchers can change, add, or subtract their interpretations of the data and rerun the analysis. Such transparency can also enhance the level of collaboration Dr. Bellin hopes for. Using tools that are freely available further enhances that ability.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-215, 2016.