Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-215-AC2, 2016

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

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: 6 June 2016

We are pleased to have the detailed review from Dr. Olaf Cirpka of our opinion piece. Dr. Cirpka raises some important issues. There are a couple issues that warrant revision of our text to clarify our intent and meaning, while there are a few others we do not agree with. We have distilled Dr. Cirpka's extensive comments to a few salient issues which we address in turn.

1. *"The authors take a recent scandal ... as an opportunity to call for equally strict rules" requiring scripting of data analysis and modeling in Hydrology as was done by the Institute of Medicine:* Indeed, we suggest that repeatability is as much an

Printer-friendly version

Discussion paper

issue in hydrology, but we do not suggest to define a set of rules that need to be strictly enforced. Rather, we suggest that the techniques mandated by IOM have relevance to hydrological science (and many other fields).

2. *"discuss counter-arguments and obstacles":* As an Opinion Piece, we prefer to rely on the online Interactive Discussion (such as this one) to form the other side of the discussion.

3. *"rethink whether repeatability is really more important than reproducibility":* This is an important topic and we regret that Dr. Cirpka interpreted our brevity on reproducibility as dismissal of its importance. We will expand the paragraph distinguishing repeatability from reproducibility to better explain why our focus is soundly on repeatability in this piece. Most hydrological studies are not fully controlled experiments, as in many other fields, but take place in the field, which means that results strongly depend on the specific field site and the specific circumstances (temperature, rainfall, river discharge, etc.), which cannot be fully reproduced. Only a few hydrological experiments are truly reproducible in the classical sense, for example at the Borden and MADE sites mentioned by the reviewer. In contrast, many hydrologists are trying to understand and predict natural systems through measurement and modeling rather than performing controlled experiments that can be reproduced by colleagues.

4. *"broaden their perspective to reach out to hydrological researchers who are not modelers":* The paper extensively discusses both data analysis and modeling, and our opinion equally holds for projects that do data analysis without modeling. But, in a sense, we are all modelers. Experimental campaigns are commonly followed by data analysis, which includes some kind of modeling, even if it is "just" trying to determine a trend.

5. *"separate issues of transparency from the call of open-source programing":* We are (as our opinion) advocates of open-source programming. However, our focus

Printer-friendly version

Discussion paper

is meant to be on the techniques of recording steps taken in analysis than on open-source software particularly. But, as we state in the paper, the two are related: "Without the availability of an executable code, the simulations can still not be repeated and without the availability of the code itself, the computational steps in the code cannot be understood and scrutinized." Nonetheless, we have tried to make it clearer that open-source programming is a factor but not the entire answer.

6. *"it gives the erroneous impression that omics was specific to cancer research":* We regret that Dr. Cirpka got this impression, but we feel the context is clear and indeed it was omics that was at issue in the cancer scandal. We propose to highlight that we are speaking of omics in the context of cancer research by adding words in bold type in the sentence "The fields of omics **as used in cancer research** and hydrology may seem as completely unrelated..." which we assume was the source of Dr. Cirpka's objection.

7. *Objection to the analogy between hydrology and life science on two accords. a) physical scientists have more connection to their raw data so the repeatability steps may be applied in hydrology, and b) there is less pressure to twist results in hydrology as lives are not at stake in the same way they are in pharmaceutical research.* On the first accord, whether hydrologists have more connection to their raw data is a judgement call that we are not able to make, but we agree that repeatability steps are equally valid for hydrology (and other sciences). On the second accord, we often hear similar arguments that hydrologic findings have less at stake (at least in the short term) than national security, health, etc. While this is true, we don't agree that such a case serves as an excuse for us (as hydrologists) to be less robust in conducting our science. If we wish to take ourselves seriously (and be taken seriously by others) we need to hold ourselves to a high standard. However, and harkening back to point 1, we are not advocating for strictly enforced rules, but rather suggesting best practices that, ones adopted

by the community, will form a standard that others want to comply with as best practices.

8. *"I agree with the authors that transparency in data processing is mandatory. But this may be achieved by other measures than enforcing everybody to use open-source codes written in a free-of-charge scripting language.":* We will revise our language to make it clear that free-of-charge is not the principal criterion we are advocating, although it makes it obviously a lot easier to repeat a modeling effort when the code is free. However, Dr. Cirpka also discusses that he is bored reading about why certain data are selected and others rejected in papers, but he adds: "...much worse would be dropping the three time series and all outliers altogether and pretending that only 12 surprisingly consistent time series were taken." We agree completely! That's why scripts are much better than spreadsheets. In a spreadsheet, often one simply deletes data that are not carried forward in analysis, but in a script that operates on data with auditable provenance, every such decision to drop a member of a dataset (or perhaps make a judgement call about data quality, such as a unit or datum conversion) can be documented.

9. *"The authors are advocates of free software. But this is not the only way of guaranteeing transparency in data processing. I don't mind excel spreadsheets, if they are well documented.":* The paragraph following this comment has a fair bit to unpack. First of all, Excel spreadsheets certainly play a role, but with complicated data analysis, often the order of operations and details about the calculations is difficult to fit properly in the confines of a spreadsheet. We advocate scripting languages in this piece, but even Fortran or C source code can contain detailed comments that might make it easier to follow the progression of calculations made to process data. RScript, MATLAB, and Python notebooks are much better than that. We also strongly disagree that placing the equations in an appendix is "absolutely sufficient." It is one thing to work out equations properly

and entirely another to actually implement them correctly. Many errors are not as egregious as using the wrong equation, but are things like a unit conversion not made (that's how to crash a spaceship into Mars!) or even a wrong sign. The scripts contain not only notes but the actual calculations and since the scripts have been written anyway, why not make them available? We totally agree that archiving data in a meaningful way is a challenge, but that's a weak excuse not to try. Forcing PhD students to be more transparent and create repeatable work would be a service to the community going forward. Finally, it's disappointing to hear Dr. Cirpka thinks that enforcing such requirements on the water-quality community is doomed to fail. Indeed, the omics research related to cancer research that motivated this piece has more in common with water-quality than with quantitative hydrology. It seems like an excuse not to try rather than a solution.

10. *"On free software":* Just three additional points on free software to address here. First, to imply that because a code has been paid for it is bug-free is naive - patches are constantly made to commercial software partly because bugs are found by users. The inability for a user to inspect the code is an impediment to quality assurance. Sure, not many users will want to look at the code, but enough do so that it can enhance quality. That said, we do not mean to imply that all software must be free. FEFLOW is a solid code, as are many in the petroleum industry which are expensive and commercial. But, it's simply not true that "If codes have been scrutinized enough by benchamrk*[sic]* tests etc. (which costs human resources, too), the users can rely on them without having access to the source code." Scrutiny and analysis by users is key to any software development and hidden bugs in proprietary code can go unnoticed much longer than when external scrutiny is ongoing. Second, open-source and free software is not the same as unmoderated community software. All three exist in various combinations, but many open-source projects are maintained by a team (who often sell training and consulting services rather than shrinkwrapped boxes of software to

Printer-friendly version

Discussion paper

pay the bills). In fact, many open-source codes can be purchased so are not technically free. Many tools have been developed for open software development to enforce rigorous testing and quality checking by a limited team. Git makes this available through putting the code online so anyone can modify their own copy, but the lead developers decide which proposed changes get accepted.

11. *Finally what the reviewer called an "ugly" comment on whether anyone will read the data analysis scripts:* It is not necessary that all the scripts be reviewed by journal editors and peer reviewers. It is fine in our view to assume that the calculations are correct while reviewing a paper. But...having such scripts and notes available to the reviewer can be valuable when results don't make sense or an error is suspected. They need not be read and scrutinized in every case to be valuable. Certainly many cases have little at stake and will not be reviewed, but in some cases the audit may be crucial.

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

Printer-friendly version

Discussion paper