

Hydrol. Earth Syst. Sci. Discuss., author comment AC3
<https://doi.org/10.5194/hess-2021-392-AC3>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC1

Caitlyn A. Hall et al.

Author comment on "A hydrologist's guide to open science" by Caitlyn A. Hall et al.,
Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-392-AC3>, 2021

- **RC1:** 'Comment on hess-2021-392', Francesca Pianosi, 02 Sep 2021 reply

The article gives an interesting contribution to the discussion on Open Science in hydrology. I especially appreciate the practical focus on helping researchers to get started with OS, and the linking to an online repository where new materials and resources will be shared beyond the article publication.

RC1.1: The manuscript needs to be revised to include the section on Principle 2, which was made available as a separate file. I do not have other substantial revisions to recommend but some points for improvement and further discussion.

Response: We are glad to hear that the reviewer found this manuscript interesting and that they found the online repository helpful. We apologize for not including Principle 2 in the original document. That was a misstep on our part and this section will be included in the revised manuscript posted on the HESS discussion board.

RC1.2: In particular, given the "practical guide" angle of this article, one suggestion could be to complement the text with Tables listing in a very concise manner the various recommendations / tips / dos & donts for each Principle. For example, for Principle 2 (open software) the list may start like:

- use open-source software such as R, python or QGIS to develop your analysis
- use open-source version control system (e.g. Git) to manage changes to your code
- include documentation as comments embedded in the code as much as possible
- etc.

I think this would help reinforce key messages and help readers navigate the (numerous) points made in the text

Response: We thank the reviewer for making this great suggestion. We agree that a concise table of tips, tools and resources for each principle would be a great way to summarize main takeaways for readers. We will include an additional table in the Practical Guide section in the revised manuscript and will add the table to the website as well.

While reading the paper I noted down several other comments. I am not sure they are all worth including/addressing in the paper but I'll report them and leave to the authors to decide if they want to take them onboard.

RC1.3: [1] L. 19 - Social challenges to embracing OS. The authors essentially mention one, the fear of being scooped, but I think others are as important. For example some researchers may be reluctant to share their software as this may bring further scrutiny and criticism of their work. Some seem to feel a sort of "jealousy" for their software, which they don't want to see modified (maybe improved!) by others. Maybe the point here is how we perceive and value intellectual ownership. If I make my software available to others so they will (unavoidably) find bugs to fix and weaknesses to improve, does this diminish or increase the value of my original contribution?

Response: Thank you, this is an important addition. We will modify the sentence to:

"...and social (e.g., fear of weaknesses being exposed or ideas being scooped) challenges remain."

We very much agree that intellectual ownership of code is a big social challenge when it comes to OS. We will include a discussion of this aspect in the revised manuscript. While we cannot solve the problem within the scope of this manuscript, our hope is that it will kick-off a wider conversation on how we, as a research community, give credit for and value open software along with its benefits for advancing hydrological research. We look to work being done by the Research Software Alliance, as they aim to do this for research broadly.

RC1.4: [2] Line 94: "sharing the entire research process and approach (e.g., failed attempts and lessons learned that impacted research outcomes)".

I totally agree although there is a tension here between conciseness (which is needed for readability) and completeness (needed for OS). I think a good way to resolve the conflict is by having unlimited "Supplementary materials" along with a paper - as some journals now allow - so that authors can keep the main article focused on key findings, while giving detailed documentation of all the research process in the SMs.

Response: We agree. We will include this as a suggestion explicitly in this line, such that it reads:

"sharing the entire research process and approach (e.g., failed attempts and lessons learned that impacted research outcomes) as appropriate in the main journal article and in more detail in the supplementary materials section of a publication. An additional option for authors is to share the entire research process associated with a publication through the Open Science Foundation's platform."

RC1.5: [3] Line 98: I think the point about "minimising the use of jargon" is very important. We use a lot of academic writing cliches in our articles, perhaps thinking it makes them sound more technically solid, but often it only makes them more difficult to read! Another issue is the recourse to hyper-specialised terms that are only understood within our small research niche - and often take different meanings across sub-communities even within the same broad discipline (a good example: the diverse uses of the term "bottom-up approach" across sub-communities in hydrology and water resource management). Every now and then initiatives are launched to build glossaries that should help researchers navigate each other jargon, but my impression is they are quickly abandoned (for example years ago I was involved in a project on uncertainty and risk in natural hazard assessment and such glossary was one of the project outputs... I don't think it was ever delivered!). Maybe rather than building glossaries we should just do more to use a common and simpler language. This should include avoiding the creation of

new terms for concepts that may be easily described with existing maths terminology. I write and review a lot of modelling/methodological papers, and I have the impression that new terms (and acronyms!) are often created under the pressure to "demonstrate novelty" - authors may be afraid that if their proposed methodology does not have a new name but is described using standard terms from a statistic textbook, reviewers will dispute its novelty. This way though we make our papers unnecessarily obscure and in the long-run we collectively contribute to fragmentation of knowledge and duplication of efforts. This links back to the general tension I find between OS and the way we reward (over-emphasise?) "novelty" of individual contributions (see also comment [1] above).

Response: Thank you to the reviewer for their insightful thoughts on this. We agree that jargon is used as a way to demonstrate novelty, and glossaries are not a solution (rather only reduce the symptoms). We will include the following a brief discussion of these two items in the text:

Jargon can be used as a way to demonstrate novelty or describe niche details, and glossaries are not a sustainable solution in the long term for supporting interdisciplinary open science progress. However, concepts can be expressed using simple fundamental terms familiar to scientists across disciplines.

RC1.6: [4] L. 193 About green OA. Maybe this sounds naive but I really wonder what are the drawbacks of "green OA" (and how it may be sustainable in the long-term for publishing companies)? Authors do not pay publication fees, readers do not pay subscription fees (after the embargo period) as they can access the non-typeset version... sounds very convenient - but for the publishing companies! Am I missing something?

Response: Green OA means that the review version of a manuscript (postprint) is shared with the public, while access to the final typeset version remains restricted to subscribers. Since the involved publishers only gain benefit from subscriptions, they have a strong incentive to make the final version more useful to readers than the postprint version. This is often achieved by formatting templates at the postprint stage that make a paper barely readable (e.g. figures separated from figure captions somewhere at the end of the document). So effectively, green-OA is not meant to work for publishers, it is a fall-back solution to tick the open-access box, while still maintaining a strong incentive to pay for the final typeset article for a more convenient reading experience. There are research communities (e.g., AI research) that mostly publish in a Diamond open access format, i.e., no fees for anyone, but obviously these do not follow a subscription-based business model, as green-OA does. We will clarify this in the main article.

RC1.7: [5] L. 226: "For software, we suggest the authors start by declaring a permissive license because it improves transparency and reduces downstream licensing conflicts."

I personally agree with this suggestion but I think many (including some at the R&D team of the University I work at!) would find it controversial. A permissive license (say for example the MIT license) implies that the software developed by publicly-funded researchers will be freely available also to users that may make a commercial use of it - and hence may be willing to pay for a licence. So, should universities "give away" a potential source of revenues? Is this fair to tax-payers who funded the research and related software development?

Response: We agree with the reviewer that a discussion on ownership of software created with public funds was lacking from our paper. We will add more discussion on this. In short, we think that software created by universities should be available for the entire public, including companies. Leaving distribution of this software in the hands of university

valorization departments severely limits adoption to the lucky few that can afford the licensing fees and time needed to negotiate access, which are usually large corporations.

Some more general thoughts:

RC1.8: [6] A lot of the activities for OS mentioned in this paper take time - for self-learning and training and to make data/software accessible to others. For example, documentation is key for open software to be meaningfully used by others but developing good documentation is very time-consuming (this was a key lesson I learnt in my own projects - see Pianosi et al 2020). So I think there is a tension here between OS and the general "publish or perish" attitude. I wonder if Open Science is necessarily also Slow Science (e.g. Frith, 2019)?

Response: We thank the reviewer for bringing up this important point about the time it takes to learn new skills, attend training, and make data/software open and accessible. We agree there is a tension between the social pressure to increase the pace of research while ensuring research is transparent and well documented. This again highlights the tension between quality and quantity, inherent in any productive environment. OS effectively shifts the balance from quantity to quality, and it necessitates new approaches of research assessment, as e.g. formulated in the SF DORA declaration (<https://sfdora.org/>). That said, we will include discussion of this balance between social pressures and open science and will include the paper on slow science that was suggested by the reviewer. We will add the importance to set up future generations of hydrologists by suggesting incorporating OS courses into curriculum, such that they are set up for success to work in an OS manner. We appreciate them pointing this paper out.

RC1.9: [7] Open teaching. I am all in favour for it - like most of us last year I developed lots of materials for on-line learning and I am keen to make it open access (as soon as I find out what's my University policy on this!). This said, I wonder what the long-term implications of open teaching will be for university life. If excellent study materials become available online to all and for all subjects, then what is the reason for enrolling in a university programme instead of self-learning? Will students only attend university to clarify doubts and get assessed - or, in other words, will the main role of universities become accreditation rather than delivering contents? I am not saying this is necessarily a bad thing (maybe it'll give academics more time for research or other type of engagement with students, such as mentoring or research-based teaching) - just highlighting it would be a very substantial change to the way higher education works today.

Response: We thank the reviewer for bringing up open education as an important part of open science. The same arguments used to argue for open science can (and should in our opinion) be used for open education. We recognize that making teaching material openly available will change the way academic knowledge is translated to students and will result in growing pains for those. As teachers ourselves we know, both from literature on education as well as from our experiences, that the roles of the teacher and peer interaction is essential in the learning process and cannot be 'completely replaced' by online available open education material. For example, networking, lab experiments, interpersonal development, etc. cannot be replicated solely by open education material. Rather, the availability of open education material offers the teacher the possibility to focus their attention on developing learning activities that reinforce concepts provided by open education material and addressing different learning styles to have the most impact on the students.

While the above answer to the reviewer shows our thoughts on Open Education, we are hesitant to add too much of an emphasis on it in the manuscript under review. The focus of the manuscript is on Open Science for hydrological researchers. We do think a separate publication (or maybe even series of publications in a special issue) on Open Education in Hydrology is timely and of great value to the hydrological teaching community. We will add the following sentence to section 1. Motivation for Open Hydrology:

We will add the following to the manuscript:

"In education, open science makes research outcomes and processes available to teachers of hydrology courses for inclusion in their teaching. A movement parallel to open science, but not the focus of this manuscript, is open education, which argues for and provides tools to share education materials and best practices freely and openly."

RC1.10: References

Pianosi et al (2020) How successfully is open-source research software adopted? Results and implications of surveying the users of a sensitivity analysis toolbox, EMS.

Available at: https://research-information.bris.ac.uk/ws/portalfiles/portal/215604556/Paper_SAFE_Survey_accepted.pdf

Frith (2019) Fast Lane to Slow Science, Trends in Cognitive Sciences.

Available at: https://discovery.ucl.ac.uk/id/eprint/10091940/1/Frith_Fast%20Lane%20to%20Slow%20Science%20Prefinal.pdf

Citation: <https://doi.org/10.5194/hess-2021-392-RC1>

Response: We will add these references to the manuscript and refer to them appropriately throughout the manuscript.