

Answer to the review of Maurits Ertzen

Reviewer: First of all, I have to apologize for being incredibly late with my review.

Then to the content of this review itself: I am afraid that my conclusion is that I think this paper is not ready yet for further review after revision. I find the topic of the paper crucial for the hydrological community, if only – but not restricted to – because the hydrological community has recently discovered that human agents matter in how water flows are shaped and that hydrological models have a role to play in the world. The socio-hydrological ideas and the HELPING initiative are signs of this, respectively, and each with some rather major issues on what they actually entail in theoretical and practical terms, if you ask me. Having said that, this paper is not the input I think we need to enhance the debate on such issues.

I have three main reasons to defend this (harsh) claim about the paper, each of which I will detail below:

1. The method is arbitrary and not transparently displayed
2. The content discussion tends to the superficial and does not engage with crucial issues
3. The conclusions are not supported by the review itself and include some strange aspects

Answer: Thank you for your comments. We will build on your suggestions to improve the manuscript, especially regarding an elaboration on the method we used, to deepen the discussion and strengthen the conclusion - including clarifying the aim and outcomes of the review. Please find our answers to your specific comments below.

Reviewer: Method

What (some of) the authors did was search for material with key words in a database and write a text. One could use specific names for this, but I do not see anything structurally different from my simple positioning of the method. That method is in itself obviously the basis for most literature reviews, but I would have liked to see much more detail about the choices – as the search itself is shaping the data (the resulting texts). The choices are described, not explained and not detailed.

Answer: We will elaborate on the methodological approach used in the literature review, specifically the ROSES method for environmental sciences (<https://www.roses-reporting.com/>), and in the thematic analysis including further explaining the choices we have made in terms of the query and the selection of texts. We have started with a review of articles that analysed elements of how models have influence, that were familiar to us as good examples, or that came up in a general search. We analysed the keywords listed in these articles to identify if a common query could be distilled, which proved to be impossible due to different jargon or different approaches to the analysis of model-use in the articles. We have then tried different queries based on keywords such as ‘influence’, ‘power’, ‘values’, ‘reflexivity’, ‘accountability’, and ‘responsibility’, or specific theories that engage with deconstructing models and their role in society, including ‘Science and Technology Studies’, or ‘Social Construction of Technology’. None of the queries yielded sets of articles that analyze the socially and ecologically differentiating effects of the use of models. This testing phase helped us to define the final query

with alternative words including 'justice', equity, politics or ethics, that resulted in the articles we reviewed in the article. We accepted that the results of the query included many articles that did not answer to our research question, based on experience with the other queries, or wish to avoid biases, as well as the interdisciplinary nature of the research, and because the query did help us to identify articles on the influence of models that were unfamiliar to us.

In the revised version of the article we will explain more about the literature review methodology and process and why the majority of the 300 articles did not qualify.

Reviewer: The step to reduce the number of texts from about 300 to about 20 suggests to me that the keyword search was actually not very useful. How somehow magically afterwards some 30 other papers were found to be added remains obscure as well. The search did not find them – other type of text perhaps, but see below on that? I am quite certain that I would come up with a different selection of texts if I would follow the method as the authors describe – assuming the method description would provide the details I would need to replicate anything.

A search strategy with the keywords runs the risk that modelling applications that may have been of interest to the review – for example how stakeholders are involved or how models are used in practices – without using the specific terms of interest of the authors of this text are not found. One can never be certain in a review to have found everything relevant, but one could do a better job to stretch the limits of available texts using approaches like citation analysis. Do these texts/authors cite each other, do they refer to similar texts – or not? These authors bring the texts together, but do the texts agree?

Answer: Thank you for the reflections on the methodology.

In terms of search strategy: We agree that the query indeed provided few results, showing the very limited number of studies on the socially and ecologically differentiating effects of the use of water models. We have used a hybrid approach to the collection of articles for three main reasons: From experience we know that articles that discuss the implementation of models in practice, or that discuss stakeholder involvement, do not necessarily discuss or reflect on the influence of models. Secondly, we wanted to avoid a bias in the collection of articles by only collecting it through general searches or only articles we were aware of. Also, the chance that papers, that are collected through a query on the topic of the influence of models, are not eligible is high; We are dealing with different jargon and words with multiple meanings (such as power), as well as a tendency to refer to the potential influence of models without elaborating or questioning it. In the article we will elaborate on this more.

We will clarify better in the revised article why the methodology was chosen, and how the papers were selected in both the narrative review and through the systematic review.

Reviewer: It would have been useful to specify much more about the papers that were finally used. Regional coverage of cases and authors, years of publication, journals, key words, type of water that is modelled, type of model used, perhaps even which equations were used – all these seemingly administrative items can already be revealing about what the community of authors does represent.

Answer: That is a good suggestion. We will include more information about the papers in a table in the supplementary material.

Reviewer: The texts that are finally used include both papers and PhD theses. Is that actually allowed – or perhaps less strict “advisable”? It does include some issues of how much papers can include in their word count compared to a thesis, right?

Answer: We do acknowledge there are different scientific practices on whether it is acceptable to refer to PhD theses and that it is a very valid discussion. For this article we have made the decision to include them as some of them are not published as articles, yet represent in-depth research on models. We therefore do see the value of including them in the analysis, but we will emphasize in the review which texts are the PhD theses.

Reviewer: How the narrative aspects were defined (on their content see below) is obscure. The reader must trust the authors that what they have selected makes sense, including which text is used for which aspect. I happen to know a few of the texts, and my selection of themes and the link to texts would have been different. A sample of $n=1$ is weak (yes, I know) but I would have used something like text analysis in combination with the admin-indicators I mentioned earlier to make my selection process of themes transparent.

Answer: We will elaborate on how the themes and subthemes emerged from the papers. In brief, we first coded the articles using an inductive or ground-up approach (see for instance Linneberg and Korsgaard, 2019). We derived the codes from the articles based on the question of what elements in the papers related to how water models and their modelling processes have influence. Through different iterations of the codes we combined them in thirteen themes/codes that we subsequently grouped in four overarching themes.

We agree that the selection of the themes is shaped by our own reading and analysis of the articles as well as our experience in modeling and research modeling practices. In revising the paper, we will be more explicit about the selection of the themes and discuss the limits of our methodology.

Reviewer: Content

The Introduction is very general and surprisingly low on references. Section 2 does include a few more, and might be integrated into the Intro. Are some of the references in 1 and 2 perhaps also feasible inputs for the review process itself?

Answer: We have chosen for a general introduction to share why we write the article, and a more specified section that discusses our definition of a model to share with the readers our understanding of what models are. We think that discussing the reason of writing the article (intro) should not be conflated with section 2 (definition of models) as these are two distinct topics, but we can enrich the introduction and link to more existing literature.

Reviewer: The definition of what a model is to the authors is very broad. I would argue that the definition holds for almost any theory. I would actually agree that hydrological models are indeed

theoretical claims, but that does not mean I would use such a generous definition. A broad use of terminology might actually obscure that the quantification aspect of models might be quite important – including differences between types of models, the actual equations that are used, the temporal and spatial coverage and steps, etcetera. This would also allow including the notion that many hydrological signals can be mimicked in many models – up to the point that rainfall can be estimated from a groundwater model, something that was obviously not the idea of that model when designed. The realization within the hydrological community that anything can be modelled has actually created a debate within that same community, perhaps not yet in the terms of the authors, but possibly of relevance for them.

Answer: Our choice of a rather broad definition of models is motivated by the wish to allow for discussions between different disciplines. More specific definitions are possible, but we found them less helpful in our interdisciplinary analysis. We have taken this broad starting point in combination with a review of articles to derive what elements can be important when understanding and engaging with the influence of models. This also includes the quantification aspect that we discuss in section 2, and also in section 4.3.1 Naturalising and legitimizing world views through models. We agree this can be made more explicit.

Regarding the reference to the debates in the hydrological community. We realize there are many debates ongoing on modelling in general as well as on specific models. We aspire to constructively contribute to these by drawing attention to the potential power of models and the role of the commissioner, modellers, users and the potentially affected, which is little debated in scientific literature.

Reviewer: The broad definition and some of the contents mentioned (see below) would have actually allowed the authors to include discussions within the hydrological community on building links with other fields, on models as I already mentioned, and comparable more interdisciplinary awareness in the hydrological community to position their model debate. Work by the group of Dr. Van Loon comes to my mind, to name just one example. I do agree that models are worth discussing, but let's not argue that they are special in terms of supporting the powerful, offer limited world views and the like. Research efforts have the tendency to do the latter, with only a few having an answer to the former.

Answer: We recognize that in the hydrological community there has been an increased interest in building links with other fields - some of the authors of the manuscript are actively working to this end. We dedicate attention to these developments in section 2 when we discuss models. We are less aware of collaborations that focus on analyzing the socially and ecologically differentiating effects of models and modeling practices themselves. We think there is value in dedicating specific attention to the influence of water models as for instance water infrastructures - shaped by water models - have large ecological and socio-economic impacts, and water models are tools that are frequently used, are often seen as neutral, and as they can be used to support the powerful. This is especially explained in section 4.3 Modelling and real-world impact. We fully agree with you on the importance of discussing models constructively, and in our article we show the value of doing this in a power-sensitive way too. We do not aim to denounce models, but our aim is to explore the potentials of modeling to contribute to more just and equitable water distributions.

Reviewer: This observation of the relevance of other (connected) topics is also clear from the content descriptions. We read about data collection as specific issue, we read about decision on project focus, and – granted – we read about modelling decisions. But why these different aspects can be connected remains unclear. The discussion of the topics does not try to connect the topics, even when several of the texts seem to cover several of the authors' themes. The topics read like an unevenly distributed shopping list, with an unequal number of texts per topics, with quite superficial summaries of the texts, and again without any cross references between texts and topics.

Answer: We have made an explicit choice to unpack and show the different elements of how models and modelling processes can have socially and ecologically differentiating effects. Based on the articles we do see that the intersection of these elements is an important contribution to the different ways models can have or gain influence, and we will dedicate attention to the interconnection of topics in the conclusion and discussion. The topics are derived from the articles, which explains the difference in length.

Review: In the discussion, suddenly the texts that were rejected for the actual review are used to make a claim about what they discuss or even represent. I am sorry, but that is simply not allowed without a proper review of these texts in itself and providing the reader with the identity of these texts.

Answer: We will elaborate in the methods section why certain texts are not included in the final review, and we will include as annex the list of the articles that we have retrieved through the query and which have not been included in the text.

Reviewer: The conclusion

The conclusions do include quite a number of references. That is at least unusual, but I would say suggests that either the authors do not have a conclusion (we read an extended discussion) or the conclusion does not yet come strongly enough out of the review.

Answer: We will review the section to make more clear for the reader what are key outcomes of the review.

Reviewer: The call for improvement if using models reads like a wish list, without the wished being confirmed as practically possible. The ideas are also rather general and surprisingly de-linked from the aspects that the authors have highlighted in section 4. Again, here the effect of missing possible texts that do discuss interesting uses of models without the terms used by the authors may be seen. Furthermore, a text like Junier (2017) shows quite clearly how relatively good intentions shift within the modelling process – in the sense that the intentions are kept, but the model does no longer align with them. The wish list is nice, but meaningless without much more discussion about implementation in actual practices.

Answer: The call for power-sensitive water modelling is based both on the review of the articles, and on work that is done in relation to the dynamics we identified. It is not a wishlist, but recommendations for doing modelling differently. To clarify this, we will elaborate on the links

between the outcomes of the review and the call, and will also clarify the different elements of the call through examples where possible.

Reviewer: One quite strong suggestion at the end is the need to move out of disciplines. This might neglect the less silo-ish nature of the hydrological discipline than the suggestion suggests, but the suggestion also drops out of the sky. The review itself does not clearly prefer this interdisciplinary aspect, so why is it so crucial? The suggestion that the issues with models come from disciplinary focuses is not even mentioned in the review. This suggests that the authors already knew the conclusion before doing the review. In itself the call for interdisciplinarity is not strange, and even done by hydrologists, but it would still need to be related to the review.

Answer: The reason we suggest interdisciplinary collaboration, including between modellers and non-modellers, is that it offers opportunities to create constructive frictions that could be used to allow more conscious decision making in model development as well as facilitate joint learning. Commissioners, modellers and model users do have a role to play here. We suggest this based on several articles that we reviewed, in which authors show that a tunnel vision based on disciplinary world views, norms and values have specific obvious and less obvious effects. This is especially clear in the conclusion of section 4.2 ‘the influence of modeller’s choices, and 4.3 ‘Modelling and real-world impact’.

We will revise the conclusion to make sure that the suggestion for interdisciplinary cooperation is better explained based on the review. We will also better explain how disciplines and related world views, norms and values influence modelling based on the reviewed articles, especially in the conclusion of section 4.2 ‘the influence of modeller’s choices, and 4.3 ‘Modelling and real-world impact’.

Reviewer: Furthermore, I find the claim that involving social sciences would solve the issues strange and in need of much more refinement. Are all social scientists equipped for and/or interested in the same issues as the authors? I would argue this is not the case. Why can the observation that many hydrological modellers do things that may be less useful be combined with the claim that collaboration between the general communities of hydrology and social sciences will solve this? The original modellers that did un-useful things will still be member of the hydro-community, right? Why are social scientists in general in the position to teach the hydrologists in general?

Answer: In the conclusion we do call for more interdisciplinary collaboration as a way forward, however we do not suggest that social scientists are ‘in the position to teach the hydrologists’. Our suggestion for interdisciplinary collaboration, which can also be with other-than-social science disciplines, is derived from the review and from our own experience in doing modeling in interdisciplinary settings. It is an invitation for curiosity and joint learning based on the gained awareness from the review that in a modelling process many steps are taken that do have impact on the outcome but these are often not reflected on. Also from the review we learn that models do have the potential to have real-world impacts while often seen as neutral tools, and these potential real-world impacts can be better understood through interdisciplinary collaboration. We will be more explicit about this in the revised version of the paper.

Reviewer: Final remarks

The topic of the text is important, but I think that the evidence that the text brings is not convincingly presented, as I have tried to argue. I would have liked to think that revisions would have been possible to continue the process, but the combination of a weak (description of the) method and a rather unbalanced analysis makes me strongly suggest that the text as is should be rejected.

Please find my handwritten notes on the text in the pdf attached. If there are any questions about remarks (including my handwriting) I am obviously available for further exchange. Again, the topic of this paper deserves further elaboration – but the paper itself needs that elaboration first even more...

Answer: Thank you for your review. We will include your suggestions in the revised manuscript, especially in relation to elaborating on the methodology, clarifying the structure of the analysis, and strengthening the conclusion based on the review. We also remain open for further exchange.