

Thank you for your compliments, close reading of the article and your engaging questions. We have answered them in the attached document.

Question: *I'm a bit confused with your choice of words. You used the passive form of verbs when describing the modeling process: "models are developed", "models are made", and "models are represented/understood" which indicates that these actions are done by some people with ethical responsibilities. But when discussing the influence or power of models, you sometimes used the active form of verbs as if a model (a simplification of the world) could act on its own, thus, hiding the people who gave power to models. Is there a reason for this choice of words? This makes it sound like you are "placing responsibility on the "model" and not on those who use and misuse them" (K. Beven).*

Answer: The words indeed represent how the way a model is designed influence the interactions with its environment. It does not mean that the responsibility is placed on the model, hence our call for power-sensitive modelling which invites those who engage with models (incl. modellers but also funders, commissioners, users), to be mindful of this possible influence. However, our words do mean that we acknowledge that also the model and modelling process can be designed differently, with effects on how the model interacts with its environment.

Question: *Would these models have the same power if they were used by people with less political power? Would these models have the same influence on the decision-making if they were funded/commissioned by people with the same political power but with a stronger political will for equitable water governance?*

Answer: These are great questions, and the influence of a model indeed differs case by case. It is also the reason why we hope more articles will be written on this topic. You rightly point out that an important part of the power of models lies in the interactions. What we would add is that the way the models are designed does influence what these interactions look like. Through the review we learn that there are certain elements that we can take into account when deciding on whether to model or not, how to design the modelling process and the model itself. From the review we derive the following elements, although we are aware more factors play a role, such as the choice of how to present the outcomes. The elements below were deducted from the review and used as structure for the paper:

Mental models and policy projects

- *Problem framing: Exploration versus consolidation*
- *Knowing the world in specific ways*
- *Working towards different versions of the world*
- *Representation: Mental models translated into, and shaped by, categories*

The influence of modellers' choices

- *Modellers' choices matter*
- *Why choices are made: familiarity, habits, and standardisation of practices*
- *Modellers developed through interaction*

The 'real-world' impact models have

- *Naturalising and legitimising world views through models*
- *Exclusive and inclusive assessments*

Engaging with non-modellers through models

- *Connecting to and disconnecting from people and places*
- *Stakeholders confronted with different realities of modelling and measuring*
- *Representation and fairness*

- *Intent: Building in reflection on engaging with the real-world from a modellers perspective*

Question: *I'm curious how "models are commissioned/funded" in the selected studies in Section 4.3. Before the modeling process was even started, how does the decision to develop a model was made? Who made that decision and provided funding for model development? Maybe these models had some real-world impacts because they were commissioned to support decision and make an impact in the first place, unlike models that are developed to advance scientific knowledge.*

Answer: Your question strongly relates to the question of reviewer Keith Beven, on why to centralise the model in the analysis and not the frames and policy process around it. You are right that this is crucial to take into account. From the articles we also deduct that these processes influences what the model looks like, and in turn the models could be used to legitimize or strengthen such decisions. Also models that are developed to advance scientific knowledge could be commissioned to support certain decisions or causes. This comes back to our point that the influence and power of models is not necessarily a negative trait.

Question: *I don't think all hydrological models are political or equally political. I think it's important to identify when, where, and how "a simplification of the world" starts to have an influence on the world. For example, we can list certain conditions that make a model more or less political (e.g., diagnostic/explanatory vs. prognostic/predictive models, open-source vs. close-source, commissioned vs. original research ideas).*

Answer: Thank you for this reflection, and we agree that not all models are political in the same ways or with the same impacts. It differs per case, and indeed measurements can be taken to change the model or modelling process with the aim to have a different impact. It starts with a reflexivity on the potential impacts of the modelling process and model itself. This is also why we have added our call for power-sensitive modelling to the review, are compiling the database, and hope for more case studies.

Question: *I think section 5 would be clearer if we could distinguish the technical and ethical responsibilities of the different actors: modelers, model-users (do you mean the ones who use model results to make decisions?), and funders in each of the four mentioned considerations.*

Answer: Section 5 is our call for power-sensitive modelling that invites for reflexivity and action for those involved with models. We indeed explicitly do not place responsibility on the modellers solely, which is often done. It is a very good idea to further work on, and collect existing, examples of how this could be done in practice. We consider this as work for the near future.

Question: *Some questions on the methodology:*

L134: what do you mean by "stood out"? It's not clear to me how you identify these aspects.

Answer: We will clarify this in the paper. Based on the review we listed different elements of how models can influence and do influence their environment. These were grouped based on themes, and four overarching themes could be identified that could not easily be grouped together.

Question: *Figure 1: "absence of elements of reflexivity". what are those elements? how did you identify those elements in the literature?*

Answer: We identified those elements firstly based on papers we shared within the group of authors, as examples of in-depth analyses of how models gain influence. Building on this, we had a first idea of what

elements mattered, and we used this to design the query. In many articles, there is a mention on what a model could/is hoped to achieve, but not a discussion on how this happens. This was the most important reason of why to include articles or not.

Question: "pre-screened articles from other sources". what are the sources? how did you pre-screen those?

Answer: Those are the articles we collected and identified ourselves. We have elaborated on this in the methodology.

Question: "studies included in qualitative synthesis - X not done". it seems you still need to update X.

Answer: We have chosen for a narrative synthesis, and not a qualitative synthesis. To avoid confusion we will adjust the figure.

Question: The fact that you have more articles by pre-screening from other sources than articles from searching using keywords suggests that the choice of keywords was not as efficient as your "pre-screening" method. Did you find the pre-screened articles in your search results?

Answer: Indeed, we have reflected on this as well in the methodological section. We found several pre-screened articles in our search results, but certainly not as many as expected. A challenge is that the topic finds itself that the intersection of different disciplines, as well as that relevant keywords are often multiple-meaning words that greatly clutter the search results. The chosen systematic search was valuable as it brought up articles we were unfamiliar with.

Question: Some articles may be interesting:

Horton, P., Schaepli, B. and Kauzlaric, M., 2022. Why do we have so many different hydrological models? A review based on the case of Switzerland. Wiley Interdisciplinary Reviews: Water, 9(1), p.e1574.

White, D.D., Wutich, A.Y., Larson, K.L. and Lant, T., 2015. Water management decision makers' evaluations of uncertainty in a decision support system: the case of WaterSim in the Decision Theater. Journal of Environmental Planning and Management, 58(4), pp.616-630.

Bijlsma, R.M., Bots, P.W., Wolters, H.A. and Hoekstra, A.Y., 2011. An empirical analysis of stakeholders' influence on policy development: the role of uncertainty handling. Ecology and society, 16(1).

Jackson, S., 2006, May. Water models and water politics: design, deliberation, and virtual accountability. In Proceedings of the 2006 international conference on Digital government research (pp. 95-104).

Answer: Thank you for sharing this! We will review them closely and add them to the database.