# Making a case for power-sensitive water modelling: a literature review

Rozemarijn ter Horst[1,2] Rossella Alba[3], Jeroen Vos[1], Maria Rusca[4], Jonatan Godinez-Madrigal[2], Lucie Babel[5], Gert Jan Veldwisch[1], Jean-Philippe Venot[6], Bruno Bonté[7], David W. Walker[1], Tobias Krueger[3]

[1]Water Resources Management Group, Department of Environmental Sciences, Wageningen University and Research, Wageningen, 6708 PB, The Netherlands,
[2]Water Governance Department, IHE Delft Institute for Water Education, Delft, 2611 AX, The Netherlands,
[3]Geography Department and Integrative Research Institute on Transformations of Human-Environment Systems (IRI THESys), Humboldt-Universität zu Berlin, Berlin, 10099, Germany
[4]Global Development Institute, University of Manchester, Manchester, M13 9PL, United Kingdom.
[5]Department of Physical Geography, Faculty of Geosciences, Utrecht University, NL-3584 CB Utrecht, The Netherlands
[6]UMR G-EAU, IRD, University Montpellier, 34000 Montpellier, France
[7]UMR G-EAU, INRAE, University Montpellier, 34000 Montpellier, France

*Correspondence to*: Rozemarijn ter Horst (rozemarijn.terhorst@wur.nl)

**Abstract.** Models are widely used to research hydrological change and risk. Yet, the power embedded in the modelling process and outcomes are often concealed by claiming their neutrality. Our review shows that in the scientific literature relatively little attention is given to the power of models to influence development processes and outcomes in water governance. At the same time, an emerging body of work offering critical insights on the political implications of hydrological models and a nuanced understanding of their application in context has begun to flourish. Drawing on this work, we call for power-sensitive modelling which includes the following considerations: Work towards just and equitable water distributions; have a broad take on modelling beyond programming and coding; contextualise water modelling to engage with impact; be transparent on the expectations and choices made; foster accountability; democratise modelling which entails giving space to multiple knowledges, multiple stakeholders, and incorporating marginalised voices of peoples and nature in all stages of the modelling process. Our call should not be understood as a suggestion to do away with modelling altogether, but rather as an invitation to interrogate how quantitative models may help to foster transformative pathways towards more just and equitable water distributions.

# 1 Introduction

Water flows and storages are increasingly researched and governed through quantitative (hydrological, hydrodynamic, socio-hydrological, hydro-economic) models. These models are used with different purposes, including documenting water distribution, exploring causal dynamics, simulating changes, predicting future conditions and informing policy making. Far from being neutral tools, models are shaped by policy projects, institutional backgrounds, specific traditions and practices of modellers, and gendered relations and experiences (Sismondo, 1999; Knorr-Cetina, Lane, 2012; 1999; MacKenzie, 2006; Melsen et al., 2018a; Addor and Melsen, 2019). Since models are complex and the places and people that develop a model may be disconnected from the places and people that use the model, unravelling how and why a model functions, and with what influence, is complicated (Kouw, 2016). Yet, we argue in this paper that this complexity is an often-missed piece of the puzzle in model commission and development, and consciously engaging with it can help to improve the models' fit for purpose or support a modelling process that contributes towards more just and equitable water distributions.

Models are not neutral, and those who commission and develop models do have choices on whether modelling should be done, as well as how. The hydrological modelling community is well aware that any one model could have turned out differently with different assumptions, simplifications, data and if different people had developed it. An iconic example is the study by Hollaender et al. (2014), in which 10 research teams were presented with increasing amounts of data from an artificially constructed catchment in order to model runoff from rainfall, leading to results varying initially by two orders of magnitude. Reflexions about modelling as a social practice and the political consequences of models in the hydrological community have been primarily in terms of how a model could be considered fit for purpose and model adequacy, uncertainty, and subjectivity (Krueger and Alba, 2022).

Beven (2019) distinguishes two kinds of purposes: accurate representation of hydrological processes and mere forecasting of hydrological variables. The latter does not necessarily require any process understanding to develop output, for instance shown recently with the resurgence of machine learning in hydrology (Nearing et al. 2021). Yet, Beven (2019) argues that an accurate process representation is needed if models are to be used for decision making. Addor & Melsen (2019) and Melsen (2022) show that institutional factors play a greater role in modellers choosing models than model adequacy in the sense of fitness for purpose. The question of model adequacy begins to gain an overtly political connotation when Beven (2019, 2022) and Hamilton et al. (2022) consider the possibility of policy makers or stakeholders to be involved in assessing whether a model is fit for purpose. Further developing this point, we would add that the developments, including increasing model complexity, attention for uncertainty, fit for purpose and involvement of stakeholders, will bring the fore ever more clearly the political nature of models, as something to utilise and as something to challenge.

A pitfall could be that discussions remain disconnected from the context the models are used in, while this could improve the modelling practice itself. Naturally, the discussions described above take the model as starting- and end-point, as the aim is to

*[handwritten margin notes:]* relevance?

what does this sentence mean?

Now we have "commission and development". The paragraph starts with the use of models.

improve a model, but the challenge will be to step out of model-land (Thompson and Smith, 2019). Since hydrological science is inherently bound to societal needs (Lane, 2014), being more explicit about the political influence of models is relevant not only from a science studies perspective but also for hydrology as a discipline and for societies at large. The aim of this article, therefore, is two-fold. First, we research how academic literature discusses the many ways models and modelling processes

65 can gain influence, also beyond their intended reach. We start from the hypothesis that indeed there is still a limited scholarship attending to the influence of models and modelling practices. Second, we draw lessons on how to engage with this political charge of water models, and eventually how to harness the power of models for progressive transformation. We begin the article by introducing our understanding of what models are. We then describe the methodology of the study and present the findings of our analysis. Based on the results we define and call for a power-sensitive approach towards modelling, and discuss

70 possible methods to facilitate implementing this in practice.

*[handwritten margin note, top]: → why?*

*[handwritten margin note, right]: Is that needed as a start?*

*[handwritten note]: So, you want to use the political means of models to develop a certain world?*

## 2 Defining models, modelling, and their power

We are aware that there are different viewpoints on what model are, and subsequently what their influence on development processes looks like, and where accountability lies. It is therefore necessary to clarify the theoretical starting point of this article. First, for the purpose of this article, we adopt a broad definition of models to capture a wide range of modelling practices

75 and that resonates with the representational view many modellers share. This view understands models as simplifications of the world that support the processing of input in various ways, to create output that is informative about the input and process. In other words, the output is influenced by the process and the input (based on Losee, 1997). The simplifications of the world are based on ideas on how the world functions or should function, enabled or limited by technology, and sustained by particular forms of (expert) knowledges, values and understandings (Haas, 1992; MacKenzie and Wajcman, 1999; Krueger and Alba,

80 2022). An example are the different ways that water is understood, from a purely physical understanding that is often applied in hydrology, taking human influences into account that is common in socio-hydrology, or seeing a deep entanglement of people and water (see Linton, 2009; Sivapalan et al., 2012). Modelling and models are used for different purposes, including to consolidate ideas about what the world is, or to explore unknown parts thereof, for instance through prediction (Morgan and Morrison, 1999; Pielke Jr, 2003; Lane, 2014). Modelling can be done in laboratory- or applied settings, and for narrowly

85 prescribed purposes such as calculating the height of a dam, or to relate to broader questions of whether that same dam should be built, or where, or for whom. These questions have a potential impact (in the case of the dam, a very imminent one) on how modellers and model-users engage with and shape the world around them (King and Kraemer, 1993).

*[handwritten note]: → defined as?*

Second, to unpack both how power is inscribed in models and how these might gain power it is essential to place our analysis in science and technology debates about what knowledge is and how it is produced. This philosophical perspective has

90 significant implications for the way modelling is understood and conceptualised. In this perspective, the modelling process, from problem identification to the development or application of the model to the generation of new information and the

*[handwritten note, bottom]: If I understand the call correctly, our understanding of models as politically constructed should lead to models building a certain new world? I am not to sure ---*

support of (policy) decisions, is not linear, although often portrayed or designed to function as such (Macnaghten, 2020; Babel and Vinck, 2022). Different parts of the model development process can run simultaneously or feedback on each other, few processes run exactly as designed on paper, and models are not made in neutral laboratory settings void of funding, norms, 95 values and ideas of what the world is and should be.

The constructivist epistemologies we build on conceptualise scientific knowledge as historically contingent, situated, and socially constructed (Latour, 2003). Science and technology studies have long argued that, scientific knowledge is "primarily as a human product, made with locally situated cultural and material resources, rather than as simply the revelation of a pre-given order of nature" (Golinski, 2005: p. xvii). In contrast to mainstream interpretation of science as neutral and objective, 100 science and technology studies conceptualise environmental knowledge as political and shaped by power relations, which determine what knowledge claims are considered more relevant and usable, how and where research should be published and, in turn, what criteria and norms scientists need to conform with (Demerritt, 2001 and 2006; Law, 2004; Stengers, 2018; Turner, 2011; Zwarteveen et al., 2017). Thus, power is an inevitable component of any piece of scientific investigation. To dedicate attention to what is seen, and how, can be illustrated by the different disciplinary, ontological and epistemological perspectives 105 of socio-hydrology and hydrosocial research (Wesselink et al., 2017). While socio-hydrology takes hydrology as starting point and adds social components to improve its representation of complex social dynamics (Lane, 2014), hydrosociology takes sociology and the complex interactions between values, significance, power relations as a starting point to explain how water and society interact. An example of this different way of thinking is the hydrosocial cycle in which water is depicted to be able to flow upstream, for instance driven by economic incentives (Linton and Budds, 2014).

110 All models, including the 'purely' physical science-based and quantitative ones, are shaped by people and their norms, values and institutions, and the models shape these in return (Bijker, 2017; Bijker et al., 1987; Latour, 2000; Latour and Woolgar, 1986; MacKenzie and Wajcman, 1999; Krueger and Alba, 2022; Saltelli and Di Fiore, 2023). This societal influence is clearest and most direct through the visual output of models, such as graphs and maps, used in decision making processes. However, there are many clearly recognizable or more hidden ways in which models also interact with social processes. It may be that 115 specific elements of the modelling process have more influence than the final product (Lane et al., 2013), for instance by (re-)producing or challenging discourses, either more or less implicitly (Krueger and Alba, 2022). In this process it matters whose information and knowledge is taken into account, who and what is represented in the process, and how. Information and knowledge enter and exit models at every stage of the development process, so the relation of models with social processes happens throughout the model development chain. Yet, it is important not to essentialize the influence of models in society, 120 and to recognise that their influence might vary from case to case. As Woolgar and Cooper (1999: p 443) argue on technology more broadly, "technology is good and bad; it is enabling and it is oppressive; it works and it does not; and, as just part of all this, it does and does not have politics".

This show two different approaches, but does not clearly show the claim that only one of the two includes power...

Our constructivist theoretical approach and broad definition of models and modelling processes help to make visible that modelling is a process that is susceptible to outside influences and in which different choices are made that shape the process and output (Demeritt, 2006; Lane, 2012). Based on the above, we argue that analysing the potential influence of models requires engaging with questions on why modelling is chosen as method to produce information, what assumptions are included in the problematization phase as well as in the data and model that is used, how available technology enables or excludes, and how the process and output are communicated and questioned, and by whom. The articles that are included in the analysis do not necessarily apply a constructivist approach, but they do discuss one or all of the aforementioned aspects.

## 3 Methodology

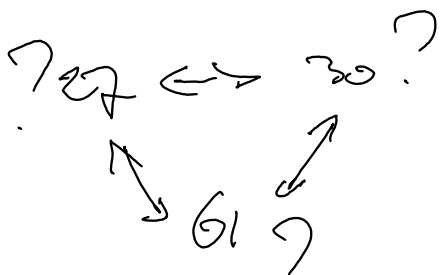
This article is based on a literature review that combines a narrative review (Cronin et al., 2008) and a systematic review to explore if and how in scientific articles is reflected on how water models have or gain influence. The narrative review was used to support the development of the query for the systematic review, and to include relevant articles that may fall outside of the scope of the query. This includes articles that were suggested by the HESS community and reviewers after the open review of this paper. The systematic part of the review is used to identify a larger set of relevant and representative articles and to examine the articles that discuss the influence of water models. We followed the ROSES (RepOrting standards for Systematic Evidence Syntheses) method (Haddaway et al., 2018), which is specifically designed for the field of environmental management. The method provides a three-staged approach that includes searching, screening and critical appraisal. It explicitly allows for additional articles to be included in the screening process to identify the best possible sample of relevant literature. This last step allows a merger with the articles obtained through the narrative review method. Each step is documented and available as additional material to this article.

In order to define the query for the systematic literature review, we first selected articles that represented engagement with the different ways in which models and technology gain or have (potential) influence, resulting in 136 articles of which 60 discuss water models, of which 26 we found are reflexive on the power the models have. For this selection we built on a diverse set of expertise as an interdisciplinary group of academics who are part of a research-collaboration on critical research on water modelling. The keywords and words used in the titles and abstracts of the articles included in this first selection were compiled, but their diversity did not form a solid basis for a query. We then turned to identifying words that related to the act of reflecting on influence and power. The final query is defined as TITLE-ABS-KEY ("water model*" OR "hydr* model*" OR "groundwater model*") AND TITLE-ABS-KEY (justice OR equit* OR politic* OR ethic*). 'Politic*' and 'equit*' were chosen as keywords because they broadly relate to how models influence issues of distribution, in relation to who gets what, when and how (Lasswell, 1936). 'Justice' and 'ethic*' were chosen to capture those articles that reflect on why certain actors – including nature – receive or are deprived of water. The query necessarily excludes words such as 'influence', 'power', 'values', 'reflexivity', 'accountability', and 'responsibility'. Earlier attempts to define a suitable query included these keywords

resulted in large quantities of articles that did not reflect on the influence models have due to the multiple meanings of these words.

Results were taken from SCOPUS and Web of Science, based on English language literature for the period January 1993 – December 2023. The query resulted in 408 unique documents. Following the ROSES protocol, we screened the articles to identify those that explicitly addressed or analysed the (potential) influence of models, including a reflection on why modelling was chosen, what assumptions are embedded in the choice for a modelling process, the input, as well as the model, how available technology enables or excludes, and how the process and output are being communicated and questioned, and by whom. We did a first screening by title and thus excluded 40 documents that had no author listed, were not in English, or did not discuss water or water models. 368 Articles were screened by abstract of which 98 abstracts showed that the article may reflect on the influence of water models and which subsequently were selected for screening the full text. Of the 98 articles, 27 articles were finally selected through the query. This included articles that applied different methods to explore the potential influence of models, including the application of multiple methods for knowledge development, participatory modelling, or actively including vulnerable groups in the analysis, but that did not explicitly reflect on power relations between the model and its environment. In relation to the narrative review, we had pre-selected 30 articles and added four suggested by the HESS community, which we included for the critical appraisal stages following the ROSES method, after the elimination of one duplicate. Figure 1 provides a graphic overview of the systematic literature review process and Appendix an overview of the 61 articles included in the literature review. Those marked with "*" were pre-selected through the narrative literature review.

*[Handwritten margin notes:]*

27 ⟷ 30 ?
→ 61 ?

Now I am confused again.

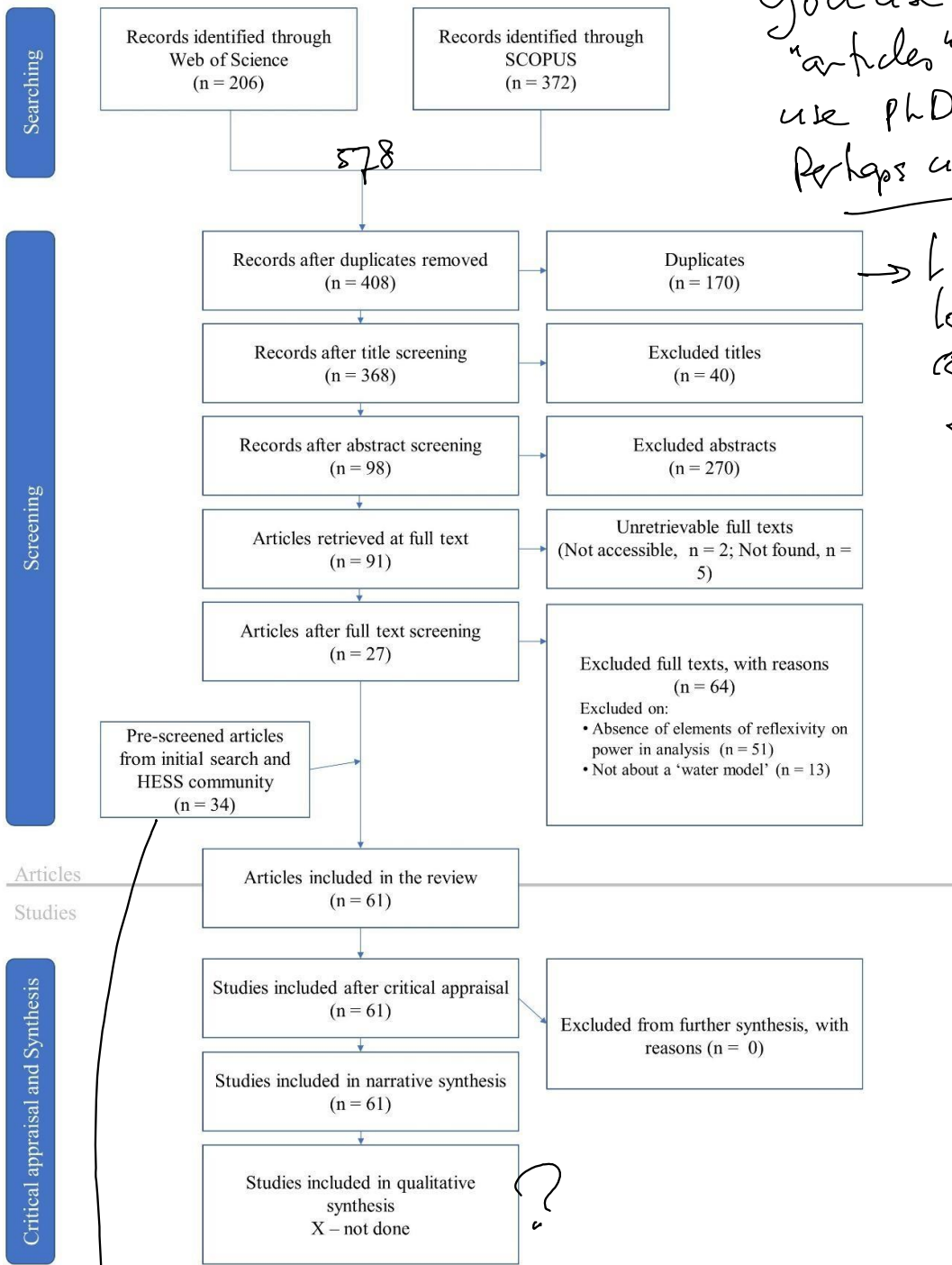
| do not clearly understand how the 34 "original" papers can be kept seperate from your new search. you did not find the 34 papers in the search?

which is a power-based decision!

you do not mention the difference between types of text in your method. I think this should be included.

**ROSES Flow Diagram for Systematic Reviews. Version 1.0**

**Figure 1: The result of the ROSES systematic literature review process**

*Handwritten annotations:*

You use the term "articles" but still use PhD theses too. Perhaps use "texts"?

→ I find this location confusing... To the innocent reader it may appear that you removed 170 from 408, but you removed 170 to arrive at 408, right?

From outer space.

So, you found 27 additional papers on top of your 34? (why) did you not find the 34 in your query? Or did you?

As the first step of the critical appraisal, the main points made in each of the articles on how models gain and have influence, or differently said, have socially and ecologically differentiating effects were identified and compared as to analyse the relationships within and between studies. Through different iterations of these central points, we finally distinguished 13 main topics that represent different phases of the production process of models, and based on these we identified four overarching themes. The main themes and related topics are:

- **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
  - Familiarity, habits, standardisation of practices and technological requirements
  - Modelling developed through interactions and institutional interests
- **The 'real-world' impact models have**
  - Naturalising and legitimising world views through models
  - Exclusive and inclusive assessments
  - The influence of presentation: colours, maps, and graphs
- **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

*[handwritten margin note: Citation analysis has not even be considered?]*

The four themes are closely interrelated and are not linear in the modelling process; thus, the fourteen topics could be seen as different layers of how models have or gain influence. They form the structure of the narrative synthesis, in which we elaborate how each theme and topic plays out in practice.

*[handwritten note: layers suggest a structure, whereas you argue it is not so ---]*

**4 Results: narrative synthesis**

This review identifies four interrelated dimensions of the modelling process that explain how models gain influence: (a) mental models and policy projects; (b) the influence of modellers' choices; (c) the 'real-world' impact models ; (d) engagement with non-modellers through models (Table 1). We present the main argument of each article reviewed under one of these four dimensions , while being aware that several articles present more than one argument. Appendix A provides more details on the articles reviewed, including the different topics discussed, as well as information on the models and case studies discussed in the articles.

| Main themes | Publication (only short reference) |
| --- | --- |

*[handwritten note: You have to structure your story, no issue there. But that does not mean your items should be seen as something structural, like layers. ~]*

| Theme | Articles |
|---|---|
| Mental models and policy projects | Alam et al., 2022; Bouleau, 2014; Budds, 2009; Deitrick et al., 2021; Fernandez, 2014; Godinez-Madrigal et al., 2019; Haeffner et al., 2018; Haeffner et al., 2021; Harvey and Chrisman, 1998; Jackson, 2006; Kroepsch, 2018; Krueger and Alba, 2022; Ländstrom et al., 2011; Lane et al., 2011; Meenar et al., 2018; Munk, 2010; Packett et al., 2020; Ramsey, 2009; Sanz et al. 2019; Shrader-Frechette, 1997; Trombley, 2017; Wesselink et al., 2017; Whatmore and Landström, 2010; Wheeler et al., 2018, |
| The influence of modellers' choices | Abbott and Vojinovic, 2014; Addor and Melsen, 2019; Alam et al., 2022; Babel et al., 2019; Bergstrom, 1991; Budds, 2009; Clark, 1998; de Oliveira Ferreira Silva, 2022; Dobson et al., 2019; Godinez-Madrigal et al., 2019; Haeffner et al., 2021; Haines, 2019; Hasala et al., 2020; Holländer et al., 2014; Jackson, 2006; Jenkins and McCauley, 2006; Junier, 2017; Kouw, 2016; Krueger and Alba, 2022; Ländstrom et al., 2011; Lane et al., 2011; Lane et al., 2013; Lane, 2014; Meenar et al., 2018; Melsen, 2022; Melsen et al., 2018; Melsen et al., 2019; Mendoza et al., 2016; Packett et al., 2020; Rainwater et al., 2005; Sanz et al., 2019; Shrader-Frechette, 1997; Srinivasan et al., 2018; Trombley, 2017; Wesselink et al., 2009; Wesselink et al; 2017; Whatmore and Landström, 2010; |
| The 'real-world' impact of models | Bouleau, 2014; Budds, 2009; Cornejo P. and Niewöhner, 2021; de Oliveira Ferreira Silva, 2022; Fernandez, 2014; Godinez-Madrigal et al., 2019; Hasala et al., 2020; Holifield, 2009; Jackson, 2006; Jensen, 2020; Kouw, 2017; Kroepsch, 2018; Krueger and Alba, 2022; Lane, 2011; Meenar et al. 2018; Melsen et al., 2018; Rainwater et al., 2005; Sanz., et al. 2019; Shrader-Frechette, 1997; Wardropper et al., 2017 |
| Engagement with non-modellers through models | Andersson, 2004; Bremer et al., 2020; Budds, 2009; Constanza and Ruth, 1998; Cornejo and Niewöhner, 2021; Falconi and Palmer, 2017; Garcia-Cuerva et al., 2016; Godinez-Madrigal et al., 2019; Haeffner et al., 2018; Holifield, 2009; Jensen, 2020; Kouw, 2017; Ländstrom et al., 2011; Lane et al., 2011; Melsen et al., 2018; Opitz-Stapleton and MacClune, 2012; Rainwater et al., 2005; Sanz et al., 2019 Wardropper et al., 2017; Wesselink et al., 2009; Wheeler et al., 2018; Wheeler et al., 2018; |

**Table 1: Overview of articles reviewed and related theme.**

## 4.1 Mental models and policy projects

210 We start with discussing the mental model (also called conceptual or perceptual model, Beven 2009; or mental images, Beck, 1999 or framing, see Odoni, N. and Lane, 2010) that underlies any numerical model. Depending on the process, the mental model is not, or less, influenced by limitations posed by data and technology and is more of an 'ideal type' than an actual model, though Krueger et al. (2016) argue that technological possibilities of what can be modelled may already co-shape what can be imagined. We divide the mental model into two elements, with the first being the ideas of how the world works, 215 including any (causal) relations, and the second being the ideas of what this world should look like. Both elements are based

on values, norms and ideas about what is important and valid to a society in general and a modelling community in particular (Haas, 1992; Haraway, 1991; Jasanoff and Kim, 2015; Morgan and Morrison, 1999). Mental models are developed based on a multitude of factors, including the common interests, backgrounds, knowledge and skills of those involved. Different communities may have very different ideas of how the world functions (Knorr-Cetina, 1999; Rusca and Di Baldassarre, 2019), or have experience with a particular way of conceptualising linked to an already familiar technology (Addor and Melsen, 2019; Babel et al., 2019; Melsen, 2022). In our systematic literature review, 22 articles dedicated specific attention to mental models. We discuss the main themes, illustrated with examples from the articles reviewed, including 1) problem framing, 2) how different ways of knowing the world influence modelling, 3) how different socio-technical imaginaries influence why a model is made, and 4) how data and categories embody world views and influence what is included and excluded and in what ways.

### 4.1.1 Problem framing: Exploration versus consolidation

Broadly speaking, there are two very distinct ways to use models. They can be used to explore unknowns, or used to consolidate ideas about reality (Morgan and Morrison, 1999; Pielke Jr, 2003). Several articles put forward how stakeholders that are part of the modelling process may have very different ideas on how the modelling process and outcomes should be used. These articles show that consolidation is often used for decision making processes in which decision makers seek to reduce uncertainty, while exploration is used in processes in which there is disagreement about the issue at hand. We use the article of Ramsey (2009) to highlight how world views, policy projects and technology intertwine based on a case study in which a GIS surface water model was created with the hope of "generating shared understandings" among stakeholders as a key strategy in reducing water allocation conflicts in the Thousand Springs Area in Idaho (USA) (p. 1975-1976). The latter objective led the modellers to try to create a scientifically sound representation of the Thousand Springs Area based on objective and measurable evidence. The model excluded some insights from inhabitants concerning the use of spring water as little measurable data was available on this issue, and the surface water model excluded groundwater from the discussions on water allocation. The exclusion of the experience of spring water users and groundwater prevented a deep exploration of the issues at hand, while this was clearly needed in the process of conflict reduction. The conclusion of the author is to call for dedicated time for exploring 'diverse problem understandings', which entails clearly defining the mental model and modelling vision, before engaging with a modelling effort.

To avoid disconnects between the model and user such as described by Ramsey (2009), Trombley (2017) suggests a multi-model approach to avoid that a model serves one particular policy project at the neglect of others. One of the suggestions they make is to design models for decision making with the aim of facilitating exploration; models becoming mediators that foster a diversity of perspectives. Constanza and Ruth (1998) propose to both engage with the consolidating and exploratory functionality that models can have in the same modelling process by introducing a three-phased modelling approach. The first stage focusses on developing the model structure and 'functional connections between variables' in discussion with stakeholders, the second stage focusses on replicating dynamics of interest realistically, and the third stage focuses on scenarios

and management options. Alam et al. (2022) propose a similar approach by calling for an inclusion of positive and negative externalities, specifically in relation to Agent Based Modelling applied to understand the impact of agricultural water management interventions. They propose such an approach as their review shows that there is limited attention for the spatially explicit and inequitable outcomes of interventions.

### 4.1.2 Knowing the world in specific ways

In the water sector, the way models are developed is often highly influenced by specific 'epistemic communities' that are bound by shared ideas on validity and causality and a way of working that engenders a particular vision of the world (Haas, 1992) or a particular way of doing through communities of practice (Lane, 2012. Bouleau (2014) shows how expertise mixes with political priorities to influence the choice of tools and issues to be addressed, and how this in turn influences the world. In the article Bouleau contrasts the approaches of two different epistemic communities in two different river basins in France. In the Rhône basin, model development was initially mainly guided by geographers and ecologists who focused on the floodplains. As a result, water was conceptualised as a 'hydrosystem' linking hydrological and ecological processes in the river and floodplains. During the same time period in the Seine basin, model development was led by engineers who assessed water quality in relation to economic development of Paris. Water was conceptualised as a condition for economic development that should be closely monitored and modelled. The mental models, differently developed based on different expertise and political priorities on top of the material properties of the two river basins, influenced what was seen and how, and consequently what the aquatic environment looked like (ibid: pp. 253). Another example is provided by Andersson (2004) who confronts a project in which three models (HBV-N, STANK, and SOIL-N) were used to assess options for reducing riverine nitrogen loads in the Upper Svarta Valley in Sweden with opinions of users. The focus of the project on nitrogen, and not on phosphorus as well, for example, was found to be limiting and not reflecting decisions that had to be taken. Despite this limited focus, the overall modelling process was deemed to create a mutual learning environment for modellers, stakeholders and decision makers. A more philosophical reflection is provided by Laborde (2015) who compares their conceptualisation of a lake through MATLAB with the conceptualisation of the same lake by a fisherman. By reflecting deeply on the underlying experiences and expertise that shape a (mental) model, they raise rhetorical questions on why the modelling version of the lake is (better) represented in decision making and the fisherman's not, and whether there is space for complexity that is brought in through lived-experiences as is done by the fisherman.

### 4.1.3 Working towards different versions of the world

Sociotechnical imaginaries are visions of what the future can become, built on a notion that technology can assist in realising this envisioned future and shaped by values (Haraway, 1985; Jasanoff and Kim, 2009). Working towards a certain envisioned future is also conceptualised as 'policy projects' (Haas, 1992). Making values explicit is therefore useful in understanding what a modelling process aims to achieve. Deitrick et al. (2021) identified and visualised what ethical and epistemological

values inspired watershed modellers in the Chesapeake Bay in the USA by surveying and interviewing the modellers involved. To support modellers and those who use or are impacted by models, the authors made visible in a flowchart what kind of choices in the modelling process related to ethics and knowledge production. These choices ranged from questions of funding and model selection, over how environmental processes were to be represented, to how users engaged with the model and how the results were interpreted, while also scoping available alternatives (ibid: pp. 12). The authors call for more openness and more explicitness by modellers when communicating these choices to contribute to transparency in decision making. Rainwater et al. (2005) show how different epistemological values and policy projects influence data collection for groundwater modelling, as well as how local political borders influence how users can engage with modelling results of a shared groundwater body in Texas. Wheeler et al. (2018a, b) also emphasised the importance of making policy projects explicit, and proposed a modelling approach for highly political and conflictual contexts in which intended model-users have very different world views and intended uses of the available water. The authors used the case of the Nile to explore possible future designs and operations of the Grand Ethiopian Renaissance Dam and its relation to operation of the High Aswan Dam in Egypt. The method did not focus on optimisation necessarily, but started with identifying upstream state and downstream state preferences as well as criteria (in this case scenarios based on acceptability and no harm) that guided the modelling exercise.

### 4.1.4 Representation: Mental models translated into, and shaped by, categories

Definitions and categories are important mechanisms to translate world views into models. Building on feminist science and making gender explicit, two articles in our literature review call for more inclusive modelling. Haeffner et al. (2021) showed that available water data often disfavour women and local communities as few disaggregated data based on these categories are available. Disaggregation, which would entail collecting specific data related for instance to gender, class, and caste, can make differences and inequalities visible. When datasets are not aggregated, or for instance create biases towards male water users who are oftentimes more visible, the modelling exercises based on biased datasets inherit the same biases and knowledge gaps unless these are explicitly acknowledged and addressed. The solution that the authors see to account for the limitations of modelling is to collect data that includes a specification including race, class, and gender, and for results to always be contextualised. This means that in addition to presenting the outputs of the modelling process, the historical and cultural context of what is modelled is described too. Packett et al. (2020) emphasise that it should not only be the input into a model that should be of concern, but that a balanced gender representation should be achieved during the whole modelling process, including problem framing and conceptualisation, model construction, documentation and evaluation, and model interpretation and decision support.

Harvey and Chrisman (1998) unpacked the development of geographical information system (GIS) technology to show how this technology can work inclusively and bring different groups together, but can also work exclusively. Based on a case study on the mapping of wetlands in the USA, the authors argue that an important element that defines who and what is included or excluded is the mental model that underlies the GIS and modelling activities. Their case started with very different ideas on

what wetlands are amongst American institutions. How different these understandings can be was highlighted in a 1995 report that compared four different datasets that represent the same wetland. The datasets disagreed on more than ninety percent of the area through different purposes, procedures, sources, definitions, and logics that shaped the different inventory techniques (Shapiro, 1995: p. xiii). To address these discrepancies, one specific system (Cowardin, 1979) was chosen as a standard by the US federal government in 1997 to define wetlands. The authors warn, however, that even though a mental model is standardised to facilitate exchange, the introduction of different modes to collect data, and different approaches to analyse these can again create different interpretations of the same area. In addition, the black-boxed nature of models can obscure these different interpretations, and an effort needs to be made to understand the influence of data collection methods and of model choices.

## 4.2 The influence of modellers' choices

The following set of articles focuses on how a model is developed. Thirty four off the articles in the review explicitly discuss modeller's choices. This includes the influence of familiarity of the modellers with the models they use, habits, as well as standardisation,

### 4.2.1 Modellers' choices matter

Modellers' choices matter, as they influence both the development and output of a model. Hollaender et al. (2009) showed through a model comparison experiment that, when provided with the same data-scarce fictive watershed, ten modellers predicted essentially ten different, and some of them very different, discharge time series based on the models of their own choosing. Within the same model, choices also matter greatly. Melsen et al. (2019) systematically demonstrated the impact of modelling decisions for the case of a flood and drought event in the Swiss Thur basin, specifically for decisions on spatial resolution, spatial representation of forcing, calibration period and performance metric. Mendoza et al. (2016) showed how hydrologic modelling decisions can influence evaluations of climate change impacts. When comparing four different modelling structures and parameter estimation strategies applied to three watersheds of the Colorado River Basin, the authors show that calibration decisions may unexpectedly have more impact than the choice of model structure. Dobson et al. (2019), by comparing eight rival framings of two models of two water resource systems in the UK, show how these specific representations of the systems influenced what water management decisions were suggested by the models. The choices of system boundaries and statistical formulation of forcing generators were shown to have the greatest impact. Krueger and Alba (2022) discuss three types of models, a socio-hydrological human-flood model, an export coefficient type model, and a water security model, to showcase the interactions between modelling and policy. These case studies are used to analyse to what extent considerations of uncertainty, subjectivity and fitness for purpose have led the hydrological community to engage with the political consequences of models and the powers inscribed in those models, be they worldviews, omissions or vested

interests. The authors especially see an opportunity for both modellers and social scientists to explore and engage the political consequences of models together, in relation to model uncertainty.

### 4.2.2 Why choices are made: familiarity, habits, standardisation of practices and technological requirements

The choice of the modelling technology or model-type is of great influence on the modelling outcomes. Addor and Melsen (2019) demonstrated, based on a survey of hydrological modellers, how familiarity with a model type is a better indicator of why a model is chosen than whether it is the best fit in terms of representing natural and social dynamics, contrary to what is typically depicted in scientific articles and consultancy reports. Babel et al. (2019) demonstrate that modellers inherit modelling choices from former supervisors and colleagues. This leads to long-lasting and sometimes unquestioned habits in model construction. Jenkins and McCauley (2006) made this visible by unpacking the GIS flow direction algorithm in ESRI products ARC/INFO, ArcView, and ArcGIS, which can seemingly make wetlands disappear from maps. Without understanding why and how the GIS algorithm functions, and without confronting the model-world with the modelled-world, this could mean that decisions are made that are ignorant of what is left invisible. Fernandez (2014) shows through historic research how the development and embedding of an indicator of minimum flow requirements (MFR) is influenced by financial and institutional needs of powerful water users in the Garonne basin in France. Originally introduced in relation to water quality, the MFR indicator later becomes a stand-alone indicator in relation to river health and to define the conditions for the construction and management of hydropower dams to define sector-based water savings. This disconnect, as well as changes in decision making processes for the host institutions of the indicator, led to the indicator to become unquestioned and blackboxed.

Whatmore and Landström (2010) trace the adoption of a formula for calculating the 'velocity or surface inclination of water flowing in an open channel of given dimensions, or Manning's n, first presented in 1889. Although it is criticised as a simplification, the formula allows for simple tuning of a model that has incorporated it, as well as limits the runtime. As such, attempts to replace this formula have failed so far. These six articles show how important the element of expertise is in modelling and warn of certain blind spots, which, once models become accepted and unquestioned tools, may be accepted as the way things are done. This does not mean that modellers are generally not reflexive. Kouw (2016) shows, for the case of hydraulic engineering in the Netherlands, different ways modellers include reflexivity in their modelling practice, including finding a balance between the detail of a model and the time needed to run it, engaging with models as 'sparring partners' instead of 'truth makers', and knowing the basic structure of the model.

### 4.2.3 Modelling developed through interactions and institutional interests

Landström et al. (2011a) draw attention to a wide range of actors that influence modelling by assessing the practices of modelling flood risk, by consultants for the Environment Agency of England and Wales. The authors show how modelling processes are shaped by environmental managers, decision makers and developers, influenced by standardised modelling processes, including practices to visit the modelled field before and after a modelling exercise, as well as long-term contractual

agreements, such as the requirement to use a particular software package. The authors argue that the high level of standardisation limits the space for asking new questions and therefore recommend that the standard practices be routinely compared with new models developed by academics. In a connected paper, Lane et al. (2011) discussed how models are used for predicting floods, taking into account climate change. By unpacking the modelling process, the authors show that a primary assumption in the model was a guideline from the government that estimated peak river flows for the 2080s will increase by 20 per cent compared to 2010. Published as part of the same research project, Lane et al. (2013) show how technology has an influence on the choice for a model. The authors discuss developments from 1D/one dimensional modelling to represent water following a specific path, to 2D/two dimensional modelling in which water can be represented to flow both down and to the sides to mimic a floodplain. A specific event, such as a flood, provided a moment in which such developments and new socio-technological constellations become apparent.

Munk (2010) and Junier (2017) also make visible in their doctoral thesis how models are developed by a multitude of actors and occurrences. In their longitudinal studies based on interviews and observations, they respectively unpacked the development process of the Hydraulic Engineering Center's River Analysis System used for flood risk analysis in the UK, and the WFD (Water Framework Directive) Explorer in the Netherlands. Wesselink et al. (2009) did a similar analysis in a research article, on how models are developed in conjunction with decision making processes. They showcased that in the case of the Dutch Meuse political considerations have an unexpectedly large influence in relation to technical water expertise, especially in relation to transboundary water management.

Jackson (2006) describes in detail the process of how CalSim, a model used by the California Department of Water Resources to estimate and plan water delivery between 2001-2021, became the topic of public controversy. Developed in a sphere of trust based on similar professional expertise, it became apparent that the model was scrutinised based on different requirements in the public sphere. This necessitated changes in the modelling practice towards more open and transparent processes. Jackson calls for a broad take on modelling, not only focusing on the conceptual, mathematical, and computer-based aspects, but also the organizational, political, and broadly sociological, which could lead to decisions to "sacrifice a degree of analytic precision and granularity, but [..] gain in broader stakeholder accessibility and general analytic wieldness" (ibid: p. 8).

**4.3 Modelling and real-world impact**

Models are often discussed within the confinement of the model-land they create (Thompson and Smith, 2019), or in other words, in laboratory conditions insulated from the public and disconnected from the world that is being modelled. Whether developed in laboratory conditions, or explicitly to inform (water) governance and management, models can have several unintended impacts. In our systematic literature review, 19 articles have dedicated specific attention to modelling and real-world impact. The articles are all based on case studies and paid particular attention to examine the context in which models are produced and how the model connects with, disconnects from and influences the surrounding environment. The two main

themes highlighted in the literature concern how models are mobilised to naturalise and legitimise certain policies and worldviews, and the ways modelling processes can work to conceal or exclude some of the affected groups.

### 4.3.1 Naturalising and legitimising world views through models

Water governance processes are always contested and political, as stakeholders are likely to hold different worldviews, including contrasting visions about the way water should be managed and allocated, and whose expertise and knowledge should be valued in decision making processes (Zwarteveen et al., 2017). Models, therefore, can have the unintended consequence of legitimising one of these worldviews whilst concealing others. To illustrate, coal mining is a contested process, in which affected stakeholders might have different perceptions on the threats and potential of this development. To illustrate, Connor et al. (2008) analysed the discourses related to a local debate on the development of an opencast coal mine in Murrurundi, a town in the Upper Hunter River basin in New South Wales, Australia. Models formed an integral part of the process by supporting the narrative of both the coalmine exploiter and the government. Despite the multiple distinct perspectives ensued by this project, the models ended up legitimising the worldviews of industry and state, whilst concealing those of many affected groups valorising care of and cultural and spiritual connections to the place and water bodies. The paper thereby highlights two real-world impacts of these models. First, they contribute to policy options grounded on notions of productivity and economic development promoted by state and industry. Second, building on this first point, models also contributed to ground the debates on scientific terminology and concepts, thereby forcing groups contesting these worldviews to draw on the same language and knowledge claims. Cornejo and Niewöhner (2021) exemplified a similar dynamic in the case of mining water abstraction in Tarapacá, Chile. Based on a groundwater model that depicted an aquifer as two separate water basins, it was decided to grant a mining company water rights as it was scientifically proven that water resources would not be affected. Here too, scientific knowledge generated through modelling was prioritised over local knowledge and everyday experiences. The way the modelling process was designed prevented affected groups from questioning assumptions on future impacts of water abstraction. In addition, as the problem was framed in the scientific language generated by the model, local communities were forced to adapt to that language and generate data that speaks to the language and arguments of scientific reports. The authors conclude that in this contested process the model became a 'real' actor, aligned with the interests of private companies and the neoliberal state. Whilst this clearly shows the political nature of models, paradoxically, it is the notion that science is value neutral that makes these models such powerful actors in water-related decision making processes.

Kroepsch (2018) and Sanz et al. (2019) also discussed how groundwater models can be used to legitimise policies even if there is limited information available. Sanz et al. (2019) showed that despite intrinsic uncertainties, and against advice of the researchers who developed the model, a MODFLOW model was used by a governmental actor to legitimise boundaries drawn that determined which farmers were compensated for refraining from irrigation, and which were not. Kroepsch (2018) questioned how it was decided to optimise space for groundwater abstraction instead of limiting it, even when impacts were unknown due to a long feedback time. Based on the analysis of 10 years of groundwater modelling and governance in the

435 Northern San Juan Basin in Colorado (USA) they argued that in this project in addition to quantitative measures, the 'human values in risk-taking or precaution' should have been prominently included.

### 4.3.2 Exclusive and inclusive assessments

When modelling is presented as a neutral scientific process, a lack of attention to the context and its power-relations can have negative effects for marginalised groups in society. An example of such a 'desocialised assessment' was provided by Budds
440 (2009) in a case of the La Ligua river basin in Chile. The author questioned the extent to which a hydrogeological model, used to represent the physical diversity in the La Ligua river basin, was representative. The model was based on data mainly available for the main river and not the tributaries, with limited information on actual water use including illegal abstractions, and the modelling process included a limited assessment of the model's validity. Despite this, the model was used to define a generic policy for the additional allocation of water rights that could have led to aquifer depletion. Budds pointed out that this
445 was possible partly due to the legitimacy given to the project by external consultants whose expertise is generally held in high regard. She further argued that the model facilitated the implementation of a policy that reproduced pre-existing water inequalities in the basin. First, the allocation of the additional water rights did not take into consideration that commercial farmers were better placed to acquire them. To illustrate, obtaining legal rights for water abstraction required a lawyer and money, thereby favouring large and smaller commercial farmers over peasant farmers. Second, Budds argues that by excluding
450 knowledge claims from peasant farmers, the model did not account for the fact that the increase in groundwater abstraction by peasant farmers was an adaptive response to the increased water use for agriculture in the valley and the 1996–1997 drought. Not recognising the vulnerability of these farmers by framing their actions as illegal ultimately increased their vulnerability. The author thus concludes that the fact that the water resources agency focused solely on hydrogeological modelling allowed the Chilean state to justify water allocation decisions that reproduced 'unequal patterns of resource use' (Budds, 2009: 418).

455 Holifield (2009) describes a similar dynamic in the case of groundwater modelling to understand the extent of pollution in St Regis, Minnesota, USA. Modelling by the Champion International Corporation was challenged by a 'counter-network' of local inhabitants and scientists, that had to prove that their representation was more scientifically viable. Holifield shows that this required them to include both disinterested ''outsiders'' and interested, locally accountable insiders, and to make connections with ''bigger'' centers of power and calculation, which can multiply and amplify the locality's connections with equipment
460 and resources (ibid, pp. 371). Inspired by Holifield (2009, 2012) Meenar et al. (2018) apply an environmental justice perspective as basis to (re)develop flood-mitigation and stormwater management plans in a watershed in southeastern Pennsylvania, USA. Using the Environmental Justice dimension of just distributions, procedure and participation, and recognition as entry points, the authors supported the redrawing of floodplains in a more inclusive way, and in interaction with local inhabitants.

465    Similar dynamics were examined by Godinez-Madrigal et al. (2019) who showed how models supported top-down management of water-scarcity issues and related water allocation policies in the Lerma-Chapala Basin, Mexico. Outcomes of one modelling exercise were not accepted when they conflicted with the interest of an important actor, and a second modelling exercise excluded an important out-of-basin user which skewed the results. The decision over water allocation was eventually enforced through influence at the highest political level, the President of Mexico. Jensen (2020) also confirmed that the power

470    of high-level decision makers plays a key role. In the case of the Mekong, the author showed there is a certain saturation in knowledge developed by models, and there is a clear limitation in their impact as governments were unwilling to build on these insights. He argued that "compared with the inventive energy deployed in modelling, moreover, it can also be observed that the efforts made by modellers to make this knowledge travel are rather less creative" (ibid: pp 88). These articles show that a model does not have influence on its own.

475    The previous examples show how models can work exclusively. The following articles show how pluralising data sources and methods can help to make the excluding nature of models visible, as well as how to mitigate this. Garcia-Cuerva et al. (2016) suggest a participatory modelling method aimed at including marginalised communities in the case of identifying opportunities for stormwater control measures in Walnut Creek watershed in North Carolina (USA). Although not yet tested, the authors opt to first develop a modelled version of the Walnut Creek, and cooperated with an NGO, Partners for Environmental Justice, to

480    facilitate discussions with stakeholders 'to evaluate alternatives and to elicit preferences' (ibid, pp 43). Hasala et al. (2020) followed up on the study of Garcia-Cuerva et al. (2016) and compared the approach of collecting information through modelling with a method that relied on interviews. Specifically looking at identifying possible sites for green roofs in majority-minority neighbourhoods in relation to stormwater control measures, they reported significant differences on what roofs should be greened based on interviews of people living in the area and the model outputs. When used in conjunction, the authors

485    showed how the model could be used as a tool to bring different stakeholders together to discuss what options fit a neighbourhood best.

### 4.3.3 The influence of presentation: colours, maps and graphs

Interestingly, few articles discuss in-depth what the influence is of specific ways of presenting the modelling results through illustrations such as graphs or maps. Most refer to this in passing. For instance, Bergstrom (1991) also concludes that ethics in

490    modelling is becoming more and more important with the rising popularity of models, and does so based on a review of the development and use of the HBV and PULSE models at the Swedish Meteorological and Hydrological Institute between 1971 and 1990. On illustrations, he calls for "Multi-colour graphical presentations are very useful for illustrative purposes but they should not be used to impress or convince where the scientific foundation is weak" (ibid, pp. 134). Abbott and Vojinovic (2014) discussed illustrations as a way to connect with stakeholders aiming that stakeholders are "challenged-out to exercise

495    and develop their own inherent knowledges, imaginations and judgments, and to exercise these both independently and interactively" (ibid: pp. 528). Also Abbott and Vojinovic point towards the responsibility of the modeller, claiming that the

"quality of the character of the modeller, becomes inseparable from the quality of the model within the quality of the total production" (ibid: pp. 528-529).

## 4.4 Engaging with non-modellers through models

500     When it comes to modelling, we want to dedicate specific attention to engagement of non-modellers in modelling processes. To counter the exclusionary nature of modelling, a popular approach is to engage those affected by the processes that the models aim to examine. Methods range from taking into account the needs and positions of different stakeholders into the design of, and communication about, the model (Cash et al., 2003; Harmel et al., 2014; Bremer et al., 2020), to different forms of participatory modelling (see for instance Voinov et al., 2010; Venot et al., 2022). Yet, few of these articles discuss power-

505     differences between those involved, account for those who disengage or who and what is excluded, or are mindful of what influences the model can have on decision making processes. In the literature review, 23 of the included articles dedicated specific attention to including people and values in a modelling process. We discern different themes, including i) engagements with how models can create connections and disconnections from the people and places that are being modelled, ii) how non-modellers relate with specific world views and policy projects included in the model, iii) representing who and what is

510     modelled in just and fair ways, and lastly iv) how modellers reflect on engaging with who and what is modelled.

### 4.4.1 Connecting to and disconnecting from people and places

Lane et al. (2011a) experiment with "doing flood risk science differently" to foster connections between academics and local people for whom flooding is a 'matter of concern, and use this as basis to co-produce knowledge in non-hierarchical ways. The project and approach created a way for local knowledge to be taken into account by the responsible institutions in the case

515     of Pickering, UK, By explicitly confronting modelling results and proposed management options with experiences and opinions of local residents, it became clear that more inclusive and less invasive flood risk management options were possible. Opitz-Stapleton and MacClune (2012) reflected in a book chapter on elements that create disconnects between affected communities and the hydrological and climatological modelling that is used for community-based climate change adaptation and disaster risk reduction. Based on case studies from the edited volume, they identified a number of issues that can create

520     disconnects between the modelling activity and the community for which it is intended. One issue that plays a significant role in communities' (dis)engagement is the degree of complexity of the model. The authors warn against thinking too much from a modelling and consultant perspective instead of a community perspective, and suggest to avoid selecting a model that is overly complex and mal-adapted to situations of data-scarcity, working at scales that are beyond the ones a community is generally thinking at (usually under 10 km), overlooking politics at transboundary and national levels, and not speaking the

525     same language of the communities for whom the model is developed. They conclude that organising modelling activities meets their proposed specifications needs "a rare combination of technical skill, cultural sensitivity, political awareness, and above

all, the time to continually engage with and build relationships within the community in order to foster resilient change." (ibid: pp 208).

An often-used framework to analyse the uptake of models is provided by Cash et al. (2003). The framework analyses how a model connects with its environment, based on its acceptance by stakeholders in relation to salience (does it fit), legitimacy (is it fair), and credibility (is it believable). We explain it here as the framework is used in two of the 48 articles included in this review. Bremer et al. (2020) applied the framework to different case studies on watershed management programs in the Atlantic Forest of Brazil. Falconi and Palmer (2017) applied it to assess whether participatory computer models for water resources management in the USA, the Solomon Islands, Senegal and Zimbabwe are indeed effective participatory decision tools based on surveys. They also emphasise that a contextual analysis is first required to gain insights into who, when, how, and why-questions. Both articles highlight that models cannot meet the expectations of each stakeholder, and therefore need to be carefully embedded in decision making processes. Bremer et al. (2020) also emphasised that it is necessary to take power dynamics into account in this process. They conclude that as hydrological modelling can influence larger development projects, it is essential to critically reflect on how and by whom these will be used and to what extent they are grounded in local realities.

### 4.4.2 Stakeholders confronted with different realities of modelling and measuring

Wardropper et al. (2017) analysed how inherent uncertainty in the Soil and Water Assessment Tool (SWAT) application to the Yahara Watershed in Wisconsin (USA) influenced the development and implementation of a water quality management programme. The programme aimed to reduce phosphorus pollution; modelling was used as a tool to estimate water quality and assign needed pollution reductions to different groups, while monitoring and compliance were based on measurements. An additional challenge in the case study was that results of the policy were not directly visible, as they were most likely to be seen within a ten-year timeframe. The authors questioned how the inherent uncertainty in this approach affected people in the watershed. The authors interviewed policy makers and those who would be subjected to the new policy on how to design such a policy in situations of uncertainty. These deliberations were found to be crucial in designing a policy that was experienced both as fair and effective, although the risk remained that the resulting actions were not influential enough to reduce the pollution. Kouw (2017) also discussed inherent traits of modelling practices that can create disconnects between models and model-users, also emphasising that uncertainty is dealt with differently by engineers, decision makers and users. Subsequently, Kouw calls for more integration of social scientists in the practice of developing and using technical tools for decision making.

Landström et al. (2011b) described in detail a participatory model experiment in which modellers, social scientists, and local residents met on a bimonthly basis over a period of one year to co-produce knowledge about flood risks in Pickering in the UK, using a 'competency group' approach. This approach asked for all participants to join as individuals, not as representatives of a certain group, and for science to be produced based on questions of the group. What was important for the project was that science was disconnected from institutions that had a role in discussions on flood risks, and that scientific questions were

not defined in advance, and were open to reframing during the project. Two models were developed as a result of this collaboration; the first was intended to be the final model and ultimately served as a starting point for discussion, and second was designed based on requests and inputs of the participants, and ultimately played a key role in shaping flood management strategy in the area.

### 4.4.3 Representation and fairness

Haeffner et al. (2018) researched how perceptions and concerns of stakeholders and decision makers were represented in the management of urban water systems in urban areas in Utah, USA. First, the authors undertook a review of sociohydrological frameworks - including models – that seek to unravel the interplay between water and society. Based on this review, they argue that sociohydrological studies tend to assume that stakeholders have "roughly equal chances of experiencing, perceiving, and responding" while generally this is not the case (ibid: pp. 666). Drawing on data collected through semi-structured interviews and surveys from city council employees, public utilities, and residents, they conclude that public officials and residents do not share the same concerns about the water supply system. Whilst residents' main concerns relate to shortages and tariffs, public officials are significantly more focused on the deterioration of water supply infrastructures. They also found citizens that were most involved in decision making were also more often shown to agree with the perspectives of water system leaders, Based on these results, they conclude that models assuming that residents are well informed and having shared understandings of the water supply system might lead to an oversimplification of socio-hydrological dynamics in a given location, and that more local involvement could mitigate this.

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

There are several authors who reflect on the impact of work in their field, and subsequently call for modellers to take an explicit ethical approach (see for instance Abbott and Vojinovic, 2014; Bergstrom, 1991). Clark (1998) also points to the responsibility of the modeller, and specifically when it comes to improved resolutions in GIS applications as "seemingly omniscient but insensitive systems" (ibid, pp 833). Although it is an old article, its reflections are still valid as technology and resolutions keep improving. Besides meeting standards for data uses and processing, facilitating access for all, and auditing, Clark also points towards the responsibility of the modeller: "Have you personally asked whether what you are doing is beneficial to the business, the customer and society? You cannot transfer this responsibility to someone else" (ibid., pp. 832). Shrader-Frechette, (1997) also call for ethical rationality in hydrogeological modelling, meaning that modelling hypotheses have to be considered in the light of their "ethical goodness" or "ethical badness" for the population on site. de Oliveira Ferreira Silva (2022) calls for a similar approach to validate models and their hypotheses, especially when it comes to the impact of its use on society. Also Lane (2014) based his suggestions for principles for socio-hydrological modellers on personal experiences with hydrology. Based on a deconstruction of practices of hydrological science, Lane proposes to i) embrace conflict and

controversy in science, ii) look for extremes to test knowledge, but doing this in a way that is sensitive to the political and ethical ramifications, iii) use real-life events to think with and step out of 'model-land', and iv) co-produce knowledge with affected groups. Lane concludes that hydrologists cannot do this alone, but that it requires both social science and hydrology.

It is this discussion in which Srinivasan et al. (2016; 2018) and Melsen et al. (2018) engaged too in a discussion on how modelling should happen. Melsen et al. (2018) pointed out that models are not value-free and that they carry significant power, which raises questions about the responsibility and accountability of those making and using models. This, the authors suggest, calls for a reflexive approach to modelling, which should incorporate questions about the model's (potential) impact, who is included and excluded and why, as well as a conscious effort to include less powerful stakeholders. In line with this idea, Srinivasan et al. (2018) proposed a number of practices to improve sociohydrological modelling, including reflecting critically on model structure and functional form, teaching people to use models as a hypothesis rather than a truth, developing guidelines on how to make modelling choices explicit, soliciting input from stakeholders, and mobilising knowledge brokers or institutions to mediate between modellers and others involved. They warn that educating scientists both in social and natural sciences takes time, and that currently the academic culture does not value interdisciplinarity.

**5 Discussion**

The literature review provides an overview of the current status of research on the influence of water models. We closely reviewed a total of 61 articles through our methodology, based on the narrative review and query (TITLE-ABS-KEY ("water model*" OR "hydr* model*" OR "groundwater model*") AND TITLE-ABS-KEY (justice OR equit* OR politic* OR ethic*). The query embodies a particular way of engaging with the influence of models grounded in the idea that modelling processes are not linear and that they shape and are shaped by society in different ways. The articles that are included in the review represent a broad spectrum of theoretical and practical approaches to the influence of water models, as well as a broad range in terms of focus. Themes include mental models and policy projects, the influence of modellers' choices, 'the real-world' impact models have, and engaging with non-modellers through models. Interestingly, we saw that in the articles reviewed there is a limited attention for the influence of vested interests – including private or academic interests - on the choice of technologies used, as well as limited attention for the way output is presented. Yet, we see that the list of themes and sub-topics forms a starting point of elements to take into account when researching or engaging with the influence of water models. Examples from the articles that were reviewed, for instance show that modelling with a particular intention in mind, such as environmental justice or gender equality, does impact the way a modelling process is done (Haeffner et al., 2018; Meenar et al., 2018). It also shows that it is useful to place discussions on the fitness-for-purpose (Beven, 2019), or on salience, credibility legitimacy discussion (Cash et al., 2005), or on a post-audits in a broader and socio-political context. Attending to the influence of models brings up questions such as 'whose purpose is served?' and 'who decided what a model should do?'.

Another observation is that several of the articles that discuss the impact of a model in practice, do not specify the specific

620  modelling software used. It is clear that choices have to be made, within the limited framework of scientific articles, on what information can be conveyed, and that interactions between specific elements within a model such as a frame or specific representation of the world are prioritised over how a model is developed (see for instance Cornejo and Niewöhner, 2021; Jackson, 2006; Kroepsch, 2018). This can be related to the studies on environmental modelling analysing the effects of the modelling process that can sometimes be more important than the model itself (Étienne, 2011). However, a case can also be

625  made that it would be useful for specific modellers to have such information, for instance through additional material to the article, perhaps especially when it comes to modelling and understanding its societal impact.

Our methodology also posed specific challenges. For example, many of the words commonly used to describe the influence of models, (including reflexivity, influence, power, accountability and responsibility) proved to be multiple-meaning words also used to describe specific – yet different – processes in modelling. This made it necessary to specify the query with the

630  risk of missing relevant articles. Also, it is known that reflexivity on these political aspects of water modelling comes in many forms and often happens in formal and informal meetings (Babel and Vinck, 2022; Melsen, 2022; Kouw, 2016). This also means that modelling processes may have been informed by reflexive practices, without being mentioned in scientific articles. However, the call to address responsibility and accountability of water modellers by Abbott and Vojinovic (2014), Lane, (2014), and Melsen, Vos, and Boelens (2018) suggest and confirm that reflexivity and acting upon it, is not a common practice.

635  Although making space and taking time for reflexivity, including discussions on ethics and accountability, adds further complexity to the modelling process. Yet, it also opens new possibilities to strengthen models and their outcomes. As the review shows, specifically when articles engage with the question of the influence of models on water governance, there are important lessons to be drawn. All authors have in common is that they spent multiple years in engagement with their respective case studies, either through longer research projects or PhD projects. It therefore calls for doing slow science (Stengers, 2017),

640  in which we get out of model-land and take time to deeply engage with the context, the tools and knowledge applied, as well as with each other.  → *new element*

Lastly and interestingly, the power disparities between those involved and affected in and through modelling processes, as well as the power of models, are addressed by only a few authors in this literature review (Budds, 2009; Godinez-Madrigal et al., 2019; Haeffner et al. 2021; Harvey and Chrisman, 1998; Holifield, 2009; Connor et al.,2008; Cornejo and Niewöhner,

645  2021; Meenar et al., 2018). Few of the articles focus on those who disengage from the modelling process or who and what is excluded, and are mindful of what influences the model can have on decision making processes. Exactly those articles, and especially the case studies that describe knowledge controversies, provide opportunities to learn, and bring up questions and examples of how accountability can look like in practice. Hence, we call for a power-sensitive approach towards modelling in the water sector. We argue that this is a crucial endeavour since models are not only influenced by power, but models also

*I find this discussion rather disappointing. It is some type of reflection but hardly engages with the results.*

*No it does not!!! A call from a selected group does not confirm anything, especially when they are the authors of this manuscript...*

650  have the power to (re)produce particular longlasting social, cultural and technical configurations in the world with more or less desirable social and sustainable outcomes.

## 6 Towards power-sensitive modelling

*illustrates*

This review shows that models shape the world around them, and the world around models shape them in return in many overt and more covert ways, influencing who is seen or not, and how, and possibly also who gets what, when and how. In section 3
655  we have argued that this is both applicable to models that are developed for practical applications, as well as those that are developed in laboratory settings (King and Kraemer, 1993). Of course, a model is never solely responsible, but there are opportunities to consciously engage with the interplay of power imbalances in the setting a model is made and applied in, as well as the influences a model has, for instance through supporting or legitimising certain policy projects, by including or excluding people or things in certain ways. Approaching models as neutral tools may conceal opportunities to do modelling in
660  support of more just and equitable water distributions. We therefore call for water modellers, commissioners, funders and model-users to further understand and engage with the power of water models, from ideation to implementation, in an ethical and accountable way.

Based on the literature review , the open review discussion process and our own experiences, we identify following considerations that can guide power-sensitive modelling (refined based on Chilvers and Kearnes, 2015; Doorn, 2012; Krueger
665  et al., 2016; Lane, 2014; Saltelli et al., 2020; Venot et al., 2021; Voinov, 2014; Zwarteveen et al., 2017;) :

- **Work towards just and equitable water distributions:** The choice for and use of water models happens in a political context and has political consequences, in a world where some gain and others are overlooked or lose. We have a responsibility to work towards more just and equitable water distributions for people and nature (See for example Abbot and Vojinovic,.2014, Bergstrom, 1991; Lane, 2014). This requires consciously defining ethical and epistemic
670  values that underlie the modelling process (See for example Deitrick et al., 2021; Holifield, 2018; Meenar et al, 2018, Packett et al., 2020).

- **Have a broad take on modelling:** The modelling process stretches beyond programming and coding, and includes everything that influences model-making and is influenced by it. For instance, it includes the processes of problematisation, defining the purpose of the model, commissioning, implementation decisions based on the
675  modelling, and the co-shaping of discussions. (See for example Jackson, 2006; Junier, 2017; Kroepsch, 2018; Munk, 2010; Trombley, 2017).

Would "knowing[24] the
ontology of your process
in quite some detail"
fit as well?

- **Contextualise water modelling to engage with impact:** Adapt the development of a water model based on a thorough understanding of the interactions with the places a model is developed and applied in (See for example Clark, 1998; Lane et al., 2011)

- **Be transparent:** Increasing transparency throughout the modelling process is a way forward to make explicit and ultimately examine and attend to the multitude of interests shaping the development and use of models and their socio-economic and ecological consequences. Modellers and commissioners can play a pivotal role in fostering such transparency, for instance by explicitly stating the underlying choices, assumptions, normative commitments and expectations as well as and tracking the choices throughout the modelling process, potentially facilitated through protocols. (See for example Babel et al., 2019; Addor and Melsen, 2019; Krueger and Alba, 2022).

- **Foster accountability:** Modellers, commissioners and model-users carry an ethical obligation to take possible real-life consequences of a modelling process or use of a model into account. (See for example: Bergstrom, 1991, Lane et al., 2011, Meenar et al, 2018).

- **Democratise modelling:** Giving space to multiple knowledges, multiple stakeholders, and incorporating marginalised voices of peoples and nature in all stages of the modelling (See for example: Lane et al., 2011; Godinez-Madrigal et al., 2019; Haeffner, 2021; Holifield 2009; Jackson, 2006; Bremer et al., 2020, Voinov, 2016).

We present these six considerations as a starting point for modellers, commissioners and users to think through the potential power-laden effects of modelling processes, and to identify possibilities to alter the design of these processes or to identify alternatives. Our call should not be understood as a suggestion to do away with modelling in the water sector altogether, but as an exploration on how to improve the practice. Although the proposed approach adds further complexity to the modelling process, it also opens new possibilities to strengthen modelling processes, models, and their outcomes.

**7 Conclusion**

*why needed?*

In this article we researched how academic literature engages with the influence water models have. Driven by an hypothesis that there are few scientific articles that critically unpack water models and related modelling processes and impacts, we have conducted both a narrative and systematic literature review to assess whether our assumption is correct, and secondly to identify what lessons we can draw from existing literature. Of the 408 articles included in the systematic literature review, 27 were finally included in critical appraisal. In addition, 30 articles were added to the critical appraisal through the narrative review and four as suggested by the HESS community. The literature review reveals how models shape, and are shaped, by the social and material aspects of the world we live in.

*The conclusion is more a call for action and comes very close to the discussion.*

705     The review also provides insights in the different ways research can be done on the influence of water models. When it comes to mental models and policy projects, it becomes clear from the articles reviewed that models are shaped by values, norms, and ideas on what the world looks like, what it should look like, and how it can be known. Models may be shaped by a group that shares similar values, norms, and ideas of how the world functions (epistemic communities), or through the negotiation between groups with very different worldviews. Developing models that are explicitly grounded in contrasting knowledge

710     claims and worldviews is often argued to be an effective strategy to avoid rejection by users or societal actors with very different worldviews. Furthermore, this approach can place models in the position of a boundary object that allows different people to use it and that facilitates collaboration. The articles also show that a mental, or conceptual, model, as an 'ideal type' of the final model, is quickly influenced and restricted by limitations in technology and data availability, and it is not always identifiable what such an 'ideal type', or multiple ideal types, would look like unconstrained by technology and data.

715     Beyond the mental models and policy choices, other more technical choices influence modelling outcomes. To further understand how choices are made, several articles analysed modelling processes based on interviews and surveys or based on practice-theoretic research to understand the role of social factors in this process (ref). It is shown that familiarity and habits to a certain extent drive the choice for specific modelling approaches, and that not only the modeller themselves, but also colleagues (Melsen, 2022), as well as managers, decision makers, and developers influence the modelling process. Kouw

720     (2016) showed that reflexivity does happen during modelling processes, such as balancing detail with run-time and considering models as objects to confront rather than accepting them as truth, although opaqueness and black boxing may make this difficult once the model is finished. → So, difference betwee development and use?

        When it comes to the impact models have on decision making and shaping the 'real' world, the case studies highlighted in section 4.3 all have in common that they are rich in description of the context the modelling practices take place in. The authors

725     dedicated attention to wishes and knowledges of stakeholders, technicalities and limitations of the model and the social consequences the model has, presenting modelling as an intrinsically social process. They also criticise modelling processes that largely overlook socio-political and economic considerations and exclude alternative views on the root-causes and dynamics of, for example, water scarcity problems.

        Lastly, engaging with non-modellers is an important theme in the articles reviewed, and they show very different approaches

730     on how to engage with non-modellers in processes where modelling is central. All start from the aim of better connecting to a wide range of people, but how this should be done and with what intention differs greatly, ranging from improving the connection between model outcomes and policies, to amplifying the wishes of the population, or to increasing fair and equitable water allocation. In relation to the influence on non-modellers, suggestions are made to address the responsibility for accountability that the modeller has by Abbott and Vojinovic (2014), Lane, (2014), and Melsen, Vos, and Boelens (2018). The

735     authors of these articles draw on personal experiences to suggest how modelling processes can be improved. Instead of taking

Hey, the same ones as page 23!

the model as a starting point, they explicitly call for reflection on the intention of modelling processes, and subsequently who and what is included and excluded. This is a distinctly different approach from the 381 articles that were excluded from the critical appraisal. The main reason for exclusion is that most articles do mention a reflection on the potential impact of the model, or the intention or expectation for the model to contribute to a more equitable and just world, but these statements are mostly brief, disconnected from a specific context, and do not make explicit as to how the model can achieve these expectations. It shows that not engaging with the potential or actual influence of models is an ignorance that is persistent.

The value and diversity of the 61 articles that are included in this review, and all differently unpack and critically engage with the influence of models on water governance, demonstrate that there is a merit in adding more case studies. Studying and doing power-sensitive modelling requires a reflexive approach that is grounded and that builds on long-term collaborations and the recognition that modelling is a complex and multifaceted process. To paraphrase Thompson and Smith (2019), this requires showcasing what happens within model-land, but also stepping out of it. As such research finds itself at a crossroads, cooperation across disciplinary boundaries is essential to nurture generative reflexivity and accountability in relation to the power of models (Chilvers and Kearnes, 2015). Moreover, several of the articles reviewed show the value of decentralising models, as well as challenging or enriching modelling results with knowledge from non-modellers and especially those affected by decisions that are related to the modelling exercises (see for instance Wardropper et al., 2017; Hasala et al. 2020). Transdisciplinary research, where both certified and noncertified water experts engage and challenge each other, seems essential (Krueger et al., 2016). This is challenging and seen as a major obstacle in a professional world that does not value complexity but promotes disciplinary thinking (Melsen, Vos, and Boelens, 2018; Srinivasan et al., 2018; Rusca and Di Baldassare, 2019).

To contribute to overcoming disciplinary thinking, we have made use of the open peer-review process of the Hydrology and Earth System Sciences journal, and invited researchers and practitioners from a broad range of disciplines to think with us, share experiences and thoughts, as well as contribute articles that have been included in an updated review and in the database of articles in Appendix A. The discussions and contributions resulting from this invitation enabled joint interrogations, now incorporated within this article, on how quantitative models may help to foster transformative pathways towards more just and equitable water distributions.

*Handwritten annotations:*

*Without discussing these in detail, I maintain my position that this claim is not allowed.*

*So, non-presence is ignorance? I do simply not always write about such stuff in some (part).*

## Appendix A: Final list of 61 articles included in the review, that explicitly engage and reflect on the power of water models

30 through a general search and personal collection
27 additional articles through the systematic review
4 through the HESS community and reviewers

Articles with * are included based on the SCOPUS and Web of Science query.
Articles with ^ are included based on the HESS community and reviewers
Based on our assessment, the "X" indicates that an article discusses explicitly i) the mental models and policy projects, ii) the influence of modellers' choices, iii) the impact models have, and/or iv) engaging with non-modellers.
"x" indicates an article discusses one of the abovementioned elements, but not explicitly.

| Title | Authors | Year | Source title | i) the mental models and policy projects | ii) the influence of modellers' choices | iii) the impact models have | iv) engaging with non-modellers | Model type discussed, and area of case study |
|---|---|---|---|---|---|---|---|---|
| Number of articles that mention a certain element of modelling explicitly | | | | 22 | 34 | 19 | 23 | |
| Towards a hydroinformatics praxis in the service of social justice | Abbott, M.; Vojinovic, Z. | 2014 | Journal of Hydroinformatics | | X | x | x | General review |
| Legacy, Rather Than Adequacy, Drives the Selection of Hydrological Models | Addor, N.; Melsen, L. A. | 2019 | Water Resources Research | | X | | | A general review on hydroinformatics, no model or area defined |
| *Understanding human-water feedbacks of interventions in agricultural systems with agent based models: a review | Alam M.F.; McClain M.; Sikka A.; Pande S. | 2022 | Environmental Research Letters | X | X | x | | Hydrologiska Byråns Vattenbalansavdelning model (HBV), the Variable Infiltration Capacity model (VIC), the mesoscale Hydrological model (mHM), the TOPography-based hydrologic model (TOPMODEL), the Precipitation Runoff Modelling System (PRMS), the Génie Rural model à 4 paramètres Journaliers (GR4J), and the Sacramento soil moisture accounting model |

→ No analysis of topics etz over time
→ Not type of texts
→ Not country of authors
→ Not country of case

| Title | Authors | Year | Source title | i) the mental models and policy projects | ii) the influence of modellers' choices | iii) the impact models have | iv) engaging with non-modellers | Model type discussed, and area of case study |
|---|---|---|---|---|---|---|---|---|
| *Experiences of the use of riverine nutrient models in stakeholders dialogues | Andersson L. | 2004 | International Journal of Water Resources Development | | | | X | General review, focused on including externalities in modelling Agricultural Water Management interventions, focus on Agent Based Modelling |
| Decision-making in model construction: Unveiling habits | Babel, Lucie; Vinck, Dominique; Karssenberg, Derek | 2019 | Environmental Modelling & Software | | X | | | HBV-N, STANK, and SOIL-N, applied in theUpper Svarta Valley in Sweden |
| *Principles and confidence in hydrological modelling | Bergstrom S. | 1991 | Nordic Hydrology | | X | | x | General review, with input of European and North American modelers in a variety of disciplines within Earth and Universe sciences |
| The co-production of science and waterscapes: The case of the Seine and the Rhône Rivers, France | Bouleau, Gabrielle | 2014 | Geoforum | X | | X | | HBV and PULSE models at the Swedish Meteorological and Hydrological Institute between 1971 and 1990 |
| Who Are we Measuring and Modeling for? Supporting Multilevel Decision-Making in Watershed Management | Bremer, Leah L.; Hamel, Perrine; Ponette-González, Alexandra G.; Pompeu, Patricia V.; Saad, Sandra I.; Brauman, Kate A. | 2020 | Water Resources Research | | | | X | Models (undefined) within the PIREN Seine and PIREN Rhône project, France |
| *Contested H2O: Science, policy and politics in water resources management in Chile | Budds J. | 2009 | Geoforum | X | X | X | X | A suite of hydrologic models, such as SWAT, InVEST, and ARIES, as well as proprietary models such as HydroBID, three watershed management programs in the Atlantic Forest of Brazil |

| Title | Authors | Year | Source title | i) the mental models and policy projects | ii) the influence of modellers' choices | iii) the impact models have | iv) engaging with non-modellers | Model type discussed, and area of case study |
|---|---|---|---|---|---|---|---|---|
| *Putting water in its place: a perspective on GIS in hydrology and water management | Clark, MJ | 1998 | Hydrological Processes | | X | x | | An undefined hydrogeological model by the National Water Directorate, La Ligua basin, Chile |
| *Watercourses and Discourses: Coalmining in the Upper Hunter Valley, New South Wales | Connor, L., Higginbotham, N., Freeman, S., and Albrecht, G | 2008 | Oceania | x | | X | X | General review, no model defined, with reflection on the US and UK. |
| Using Dynamic Modeling to Scope Environmental Problems and Build Consensus | Constanza, Robert; Ruth, Matthias | 1998 | Environmental Management | x | x | | X | An undefined hydrological model used by the Bickham Coal Company, Upper Hunter Valley, New South Wales |
| *How Central Water Management Impacts Local Livelihoods: An Ethnographic Case Study of Mining Water Extraction in Tarapaca, Chile | Cornejo, SM; Niewöhner, J | 2021 | Water | x | | X | X | Undefined hydrological models, Tarapacá, Chile |
| *The Challenge of Model Validation and Its (Hydrogeo)ethical Implications for Water Security | de Oliveira Ferreira Silva C. | 2022 | Studies in Computational Intelligence | | | | | General review, related to hyrdogeological modelling |
| *Investigating the Influence of Ethical and Epistemic Values on Decisions in the Watershed Modeling Process | Deitrick A.R.; Torhan S.A.; Grady C.A. | 2021 | Water Resources Research | X | x | | x | a wide array of models, such as the Soil & Water Assessment Tool (SWAT), SPAtially Referenced Regressions on Watershed attributes (SPARROW), and Chesapeake Assessment Scenario Tool (CAST), Chesapeake Bay Watershed |
| How Important Are Model Structural and Contextual Uncertainties when Estimating the Optimized Performance | Dobson, Barnaby; Wagener, Thorsten; Pianosi, Francesca | 2019 | Water Resources Research | | | X | | Simulated Water Resources System models, South West of the UK (research on effect of framings in models) |

| Title | Authors | Year | Source title | i) the mental models and policy projects | ii) the influence of modellers' choices | iii) the impact models have | iv) engaging with non-modellers | Model type discussed, and area of case study |
|---|---|---|---|---|---|---|---|---|
| of Water Resource Systems? | | | | | | | | |
| An interdisciplinary framework for participatory modeling design and evaluation—What makes models effective participatory decision tools? | Falconi, Stefanie M.; Palmer, Richard N. | 2017 | Water Resources Research | | | | X | Shared Vision Model (System Dynamic model built on STELLA) for the Tri-State Water Conflict in the ACT-ACF River Basin, USA; System Dynamic Model, Las Vegas, Nevada; Bayesian Network; Solomon Islands |
| Much Ado About Minimum Flows…Unpacking indicators to reveal water politics | Fernandez, Sara | 2014 | Geoforum | X | x | X | x | Undefined hydraulic and hydrological models, Garonne system, France |
| *Exploring Strategies for LID Implementation in Marginalized Communities and Urbanizing Watersheds | Garcia-Cuerva L.; Berglund E.Z.; Rivers L. | 2016 | World Environmental And Water Resources Congress 2016: | | | | X | hydrologic/hydraulic stormwater modeling system d using HEC-HMS and SWMM, Walnut Creek Watershed in Raleigh, North Carolina |
| *Production of competing water knowledge in the face of water crises: Revisiting the IWRM success story of the Lerma-Chapala Basin, Mexico | Godinez-Madrigal J.; Van Cauwenbergh N.; van der Zaag P. | 2019 | Geoforum | X | X | X | X | system dynamics models, Lerma-Chapala basin, Mexico |
| *Representation justice as a research agenda for socio-hydrology and water governance | Haeffner M.; Hellman D.; Cantor A.; Ajibade I.; Oyanedel-Craver V.; Kelly M.; Schifman L.; Weasel L. | 2021 | Hydrological Sciences Journal | X | X | | | General review, for (socio)hydrological modelling |

| Title | Authors | Year | Source title | i) the mental models and policy projects | ii) the influence of modellers' choices | iii) the impact models have | iv) engaging with non-modellers | Model type discussed, and area of case study |
|---|---|---|---|---|---|---|---|---|
| *Social Position Influencing the Water Perception Gap Between Local Leaders and Constituents in a Socio-Hydrological System | Haeffner, M; Jackson-Smith, D; Flint, CG | 2018 | Water Resources Research | X | | | X | Socio-hydrological/coupled system models, WasatchRange Metropolitan Area, Northern Utah |
| Reckoning Resources: Political Lives of Anticipation in Belize's Water Sector | Haines, Sophie | 2019 | Technology Studies | | X | | x | GIS software and the N-SPECT (nonpoint-source pollution and erosion comparison tool), Belize |
| Boundary Objects and the Social Construction of GIS Technology | Harvey, F; Chrisman, N | 1998 | Environment and Planning A: Economy and Space | X | | | | GIS technology, including ATKIS standard database model and A L K / A T K I S - G I A P software, applied to wetlands in the USA |
| *Green infrastructure site selection in the Walnut Creek wetland community: A case study from southeast Raleigh, North Carolina | Hasala, D; Supak, S; Rivers, L | 2020 | Landscape And Urban Planning | x | X | X | | Participatory Geographic Information Systems, Walnut creek, southeast Raleigh, North Carolina |
| *How to speak for aquifers and people at the same time: Environmental justice and counter-network formation at a hazardous waste site | Holifield R. | 2009 | Geoforum | | | X | X | Groundwater models:SLAEM (Single-Layer Analytic Element Model), MLAEM (Multi-Layer Analytic Element Model), MODFLOW, St. Regis, Minnesota, USA |
| Impact of modellers' decisions on hydrological a priori predictions | Holländer, H. M.; Bormann, H.; Blume, T.; Buytaert, W.; Chirico, G. B.; Exbrayat, J.-F.; Gustafsson, D.; Hölzel, H.; Krauße, T.; Kraft, P.; Stoll, S.; Blöschl, G.; Flühler, H. | 2014 | Hydrology and Earth System Sciences | | X | | | DWRSIM, used by the California Department of Water Resources to manage the State Water Project; and PROSIM, used by the Bureau of Reclamation in its Central Valley operations |

| Title | Authors | Year | Source title | i) the mental models and policy projects | ii) the influence of modellers' choices | iii) the impact models have | iv) engaging with non-modellers | Model type discussed, and area of case study |
|---|---|---|---|---|---|---|---|---|
| *Water models and water politics: Design, deliberation, and virtual accountability | Jackson S. | 2006 | ACM International Conference Proceeding Series | X | X | X | x | CalSim (generalised model for reservoir analysis, FORTRAN), California, USA |
| *GIS, sinks, fill, and disappearing wetlands: Unintended consequences in algorithm development and use | Jenkins D.G.; McCauley L.A. | 2006 | Proceedings of the ACM Symposium on Applied Computing | | X | | | General review, based on ARC/INFO, ArcView, and ArcGIS, applied to wetlands |
| A flood of models: Mekong ecologies of comparison | Jensen, Casper Bruun | 2020 | Social Studies of Science | | | X | X | Different models, including MRC SWAT, MIKE, HEC ResSIM, applied to (parts of) the Mekong river |
| Modelling expertise: Experts and expertise in the implementation of the Water Framework Directive in the Netherlands | Junier, S.J. | 2017 | Delft University of Technology | x | X | | x | Water Framework Directive Explorer, the Netherlands |
| Standing on the shoulders of giants—and then looking the other way? Epistemic opacity, immersion, and modeling in hydraulic engineering | Kouw, M. | 2016 | Perspectives on Science | | X | | x | General review, Hydraulic engineering models, The Netherlands |
| Risks in the Making: The Mediating Role of Models in Water Management and Civil Engineering in the Netherlands | Kouw, Matthijs | 2017 | Berichte zur Wissenschafts geschichte | | | X | X | General review, Hydraulic engineering models, The Netherlands |
| Groundwater Modeling and Governance: Contesting and Building (Sub)Surface Worlds in Colorado's Northern San Juan Basin | Kroepsch, Adrianne C. | 2018 | Engaging Science, Technology, and Society | | | X | x | Groundwater models (by 3M project, CBM, AHA, and Questa), Northern San Juan Basin, USA |

| Title | Authors | Year | Source title | i) the mental models and policy projects | ii) the influence of modellers' choices | iii) the impact models have | iv) engaging with non-modellers | Model type discussed, and area of case study |
|---|---|---|---|---|---|---|---|---|
| ^Ontological and epistemological commitments in interdisciplinary water research: Uncertainty as an entry point for reflexion. | Krueger, T., & Alba, R. | 2022 | Frontiers in Water | X | X | X | | ELCOM, supported by MATLAB, Lake Como, Italy |
| *Environmental Research from Here and There: Numerical Modelling Labs as Heterotopias | Laborde, S. | 2015 | Environment and Planning D: Society and Space | X | x | | | socio-hydrological human-flood models, an export coefficient type model, water security model of Dadson et al. (2017) |
| Virtual engineering: computer simulation modelling for flood risk management in england | Landström, C., Whatmore, S.J., Lane, S.N., | 2011 | Science & Technology Studies | X | | | | Discussion of different models, including ISIS, HEC-RAS and MIKE11, HEC-RAS, etc. at the Environment Agency of England and Wales |
| *Coproducing flood risk knowledge: redistributing expertise in critical 'participatory modelling' | Landström, C; Whatmore, SJ; Lane, SN; Odoni, NA; Ward, N; Bradley, S | 2011 | Environment And Planning A-Economy And Space | x | x | x | X | CRUM2D v 3.1, Pickering, UK and Wales |
| ^Doing flood risk science differently: an experiment in radical scientific method. | Lane, S. N., Odoni, N., Landström, C., Whatmore, S. J., Ward, N., & Bradley, S | 2011 | Transactions of the Institute of British Geographers | x | x | X | X | FEH & ISIS's routing methodology, Pickering, UK and Wales |
| ^Explaining rapid transitions in the practice of flood risk management. | Lane, S.N., November, V., Landström, C. and Whatmore, S.J. | 2013 | Annals of the Association of American Geographers | | X | | | Flood mapping science (HEC-RAS, ISIS and MIKE-11, RMA2 TELEMAC-2D model) |
| *Acting, predicting and intervening in a socio-hydrological world | Lane, SN | 2014 | Hydrology And Earth System Sciences | | X | | x | General overview |

| Title | Authors | Year | Source title | i) the mental models and policy projects | ii) the influence of modellers' choices | iii) the impact models have | iv) engaging with non-modellers | Model type discussed, and area of case study |
|---|---|---|---|---|---|---|---|---|
| Imagining flood futures: risk assessment and management in practice | Lane, Stuart N.; Landström, Catharina; Whatmore, Sarah J. | 2011 | Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences | X | X | | | Flood Estimation Handbook based models, UK and Wales |
| *Planning for watershed-wide flood-mitigation and stormwater management using an environmental justice framework | Meenar M.; Fromuth R.; Soro M. | 2018 | Environmental Practice | X | X | X | | ArcGIS, HEC-HMS, HEC-RAS, and HEC-GeoRas software, Pennsylvania, US |
| *What is the role of the model in socio-hydrology? Discussion of "Prediction in a socio-hydrological world"* | Melsen L.A.; Vos J.; Boelens R. | 2018 | Hydrological Sciences Journal | | X | X | X | General review |
| It Takes a Village to Run a Model—The Social Practices of Hydrological Modeling | Melsen, Lieke A. | 2022 | Water Resources Research | | X | | | Hydrologic modelling, Western Europe |
| Subjective modeling decisions can significantly impact the simulation of flood and drought events | Melsen, Lieke A.; Teuling, Adriaan J.; Torfs, Paul J.J.F.; Zappa, Massimiliano; Mizukami, Naoki; Mendoza, Pablo A.; Clark, Martyn P.; Uijlenhoet, Remko | 2019 | Journal of Hydrology | | X | | | Three Variable Infiltration Capacity (VIC) models (version 4.1.2.i), Thur Basin, Switzerland |
| How do hydrologic modeling decisions affect the portrayal of climate change impacts? | Mendoza, Pablo A.; Clark, Martyn P.; Mizukami, Naoki; Gutmann, Ethan D.; Arnold, Jeffrey R.; Brekke, Levi D.; | 2016 | Hydrological Processes | | X | | | Including Weather Research and Forecasting (WRF) regional climate model, Noah-LSM,hompson mixed-phase cloud micro-physics scheme, Colorado River Basin, USA |

| Title | Authors | Year | Source title | i) the mental models and policy projects | ii) the influence of modellers' choices | iii) the impact models have | iv) engaging with non-modellers | Model type discussed, and area of case study |
|---|---|---|---|---|---|---|---|---|
| | Rajagopalan, Balaji | | | | | | | |
| Risking the flood: Cartographies of things to come | Munk | 2010 | PhD dissertation: Linacre College, University of Oxford | X | x | x | x | HEC-RAS 4.0, UK |
| *Scientific and social uncertainties in climate change: The Hindu Kush-Himalaya in regional perspective | Opitz-Stapleton S.; MacClune K. | 2012 | Community, Environment and Disaster Risk Management | x | | x | X | Different Community Based Modelling inititatives, Hindu Kush-Himalaya |
| Mainstreaming gender into water management modelling processes | Packett, Evangeline; Grigg, Nicola J.; Wu, Joyce; Cuddy, Susan M.; Wallbrink, Peter J.; Jakeman, Anthony J. | 2020 | Environmental Modelling & Software | X | X | | | Biophysical modelling guidelines, general |
| *Impact of political, scientific and non-technical issues on regional groundwater modeling: Case study from Texas, USA | Rainwater K.; Stovall J.; Frailey S.; Urban L. | 2003 | Developments in Water Science | x | X | X | X | MODFLOW based groundwater model, Texas, USA |
| GIS, modeling, and politics: On the tensions of collaborative decision support | Ramsey, Kevin | 2009 | Journal of Environmental Management | X | | x | x | GIS Surface water model, Idaho, USA |
| The social construction and consequences of groundwater modelling: insight from the Mancha Oriental aquifer, Spain | Sanz, David; Vos, Jeroen; Rambags, Femke; Hoogesteger, Jaime; Cassiraga, Eduardo; Gómez-Alday, Juan José | 2019 | International Journal of Water Resources Development | X | X | X | X | A groundwater model, Spain |

| Title | Authors | Year | Source title | i) the mental models and policy projects | ii) the influence of modellers' choices | iii) the impact models have | iv) engaging with non-modellers | Model type discussed, and area of case study |
|---|---|---|---|---|---|---|---|---|
| *Hydrogeology and framing questions having policy consequences | Shrader-Frechette K. | 1997 | Philosophy of Science | X | X | x | | USA, the Yucca Mountain in Nevada and Maxey flats, Kentucky |
| Moving socio-hydrologic modelling forward: unpacking hidden assumptions, values and model structure by engaging with stakeholders: reply to "What is the role of the model in socio-hydrology?" | Srinivasan, V.; Sanderson, M.; Garcia, M.; Konar, M.; Blöschl, G.; Sivapalan, M. | 2018 | Hydrological Sciences Journal | x | X | | | Socio-hydrological models, general overview |
| An Environmental Anthropology of Modeling | Trombley | 2017 | PhD dissertation: University of Maryland, College Park | X | X | x | x | Chesapeake Bay Modelling System, Chesapeake Bay, USA |
| Uncertain monitoring and modeling in a watershed nonpoint pollution program | Wardropper, Chloe B. , Sean Gillon, Adena R.Rissman | 2017 | Land Use Policy | | | X | X | SWAT, Wisconsin, USA |
| *Hydrology and hydraulics expertise in participatory processes for climate change adaptation in the Dutch Meuse | Wesselink A.; De Vriend H.; Barneveld H.; Krol M.; Bijker W. | 2009 | Water Science and Technology | | X | | X | WAQUA, SOBEK, Meuse Basin, The Netherlands |
| ^Manning's N - Putting roughness to work | Whatmore S.J.; Landström C. | 2010 | How Well do Facts Travel? The Dissemination of Reliable Knowledge | X | X | | | 1D floodrisk modelling, TUFLOW, general review |
| *Exploring Cooperative Transboundary River Management Strategies for the Eastern Nile Basin | Wheeler K.G.; Hall J.W.; Abdo G.M.; Dadson S.J.; Kasprzyk | 2018 | Water Resources Research | X | | | X | Eastern Nile RiverWare Model, MOEA = multiobjective evolutionary algorithm, Nile Basin |

| Title | Authors | Year | Source title | i) the mental models and policy projects | ii) the influence of modellers' choices | iii) the impact models have | iv) engaging with non-modellers | Model type discussed, and area of case study |
|---|---|---|---|---|---|---|---|---|
| | J.R.; Smith R.; Zagona E.A. | | | | | | | |
| Modelling to bridge many boundaries: the Colorado and Murray-Darling River basins | Wheeler, Kevin G.; Robinson, Catherine J.; Bark, Rosalind H. | 2018 | Regional Environmental Change | | | x | X | The Colorado River Basin in North America and the Murray-Darling Basin in southeastern Australia |

**Author contribution**:

All authors contributed to the conceptualization and to the narrative review. Rozemarijn ter Horst and Jeroen Vos developed the query for the systematic literature review. Rozemarijn ter Horst and Rossella Alba developed the methodology. Rozemarijn ter Horst did the data collection and analysis for the systematic literature review and wrote the original draft. Jeroen Vos, Rossella Alba, Maria Rusca, David W. Walker and Tobias Krueger reviewed and edited closely, and all authors reviewed.

**Competing interests**

The authors declare that they have no conflict of interest.

**Acknowledgements**

This article is the follow-up activity of the workshop 'Towards a reflexive approach: Connecting critical research on water modelling', co-organized by the Hydrology and Society group at IRI-THESys, Humboldt University of Berlin and the Water Resources Management Group of Wageningen University on 19 and 20 January 2022. We thank all the participants for sharing their experiences and reflections about modelling waters. We also thank the reviewers and members of the HESS community for their constructive comments and engagement.

# References

Abbott, M. B. and Vojinovic, Z.: Towards a hydroinformatics praxis in the service of social justice, J. Hydroinformatics, 16, 516–530, https://doi.org/10.2166/hydro.2013.198, 2014.

Addor, N. and Melsen, L. A.: Legacy, Rather Than Adequacy, Drives the Selection of Hydrological Models, Water Resour. Res., 55, 378–390, https://doi.org/10.1029/2018WR022958, 2019.

Alam, M.F., McClain, M., Sikka, A. and Pande, S.: Understanding human–water feedbacks of interventions in agricultural systems with agent based models: A review. Environ Res Lett, *17*(10), 103003, 2022.

Andersson, L.: Experiences of the use of riverine nutrient models in stakeholder dialogues, Int. J. Water Resour. Dev., 20, 399–413, https://doi.org/10.1080/0790062042000248547, 2004.

Babel, L. and Vinck, D.: The "sticky air method" in geodynamics: Modellers dealing with the constraints of numerical modelling, Rev. Anthropol. Connaiss., 16, https://doi.org/10.4000/rac.27795, 2022.

Babel, L., Vinck, D., and Karssenberg, D.: Decision-making in model construction: Unveiling habits, Environ. Model. Softw., 120, 104490, https://doi.org/10.1016/j.envsoft.2019.07.015, 2019.

Bergström, S.: Principles and confidence in hydrological modelling. Hydrol. Research, 22(2),.123-136, https://doi.org/10.2166/nh.1991.0009, 1991.

Beven, K.: Environmental modelling: an uncertain future?, CRC Press, 2009.

Beven K.: How to make advances in hydrological modelling. Hydr. Res., 50(6), 1481-1494 https://doi.org/10.2166/nh.2019.134, 2019.

Bijker, W.: Constructing Worlds: Reflections on Science, Technology and Democracy (and a Plea for Bold Modesty), Engag. Sci. Technol. Soc., 3, 315, https://doi.org/10.17351/ests2017.170, 2017.

Bijker, W. E., Hughes, T. P., and Pinch, T. (Eds.): The Social construction of technological systems: new directions in the sociology and history of technology, MIT Press, Cambridge, Mass, 405 pp., 1987.

Bouleau, G.: The co-production of science and waterscapes: The case of the Seine and the Rhône Rivers, France, Geoforum, 57, 248–257, https://doi.org/10.1016/j.geoforum.2013.01.009, 2014.

Bremer, L. L., Hamel, P., Ponette-González, A. G., Pompeu, P. V., Saad, S. I., and Brauman, K. A.: Who Are we Measuring and Modeling for? Supporting Multilevel Decision-Making in Watershed Management, Water Resour. Res., 56, https://doi.org/10.1029/2019WR026011, 2020.

Budds, J.: Power, Nature and Neoliberalism: The Political Ecology of Water in Chile, Singap. J. Trop. Geogr., 25, 322–342, https://doi.org/10.1111/j.0129-7619.2004.00189.x, 2004.

Cash, D., Clark, W. C., Alcock, F., Dickson, N., Eckley, N., and Jäger, J.: Salience, Credibility, Legitimacy and Boundaries: Linking Research, Assessment and Decision Making, SSRN Electron. J., https://doi.org/10.2139/ssrn.372280, 2003.

Clark, M.J.: Putting water in its place: A perspective on GIS in hydrology and water management. Hydrological Processes, 12(6), 823-834, https://doi.org/10.1002/(SICI)1099-1085(199805)12:6<823::AID-HYP656>3.0.CO;2-Z, 1998.

Chilvers, J. and Kearnes, M. (Eds.): Remaking participation: towards reflexive engagement, in: Remaking Participation: Science, Environment and Emergent Publics, Routledge, 28, 2015.

Connor, L., Higginbotham, N., Freeman, S., and Albrecht, G.: Watercourses and Discourses: Coalmining in the Upper Hunter Valley, New South Wales, Oceania, 78, 76–90, https://doi.org/10.1002/j.1834-4461.2008.tb00029.x, 2008.

Cornejo P., S. M. and Niewöhner, J.: How Central Water Management Impacts Local Livelihoods: An Ethnographic Case Study of Mining Water Extraction in Tarapacá, Chile, Water, 13, 3542, https://doi.org/10.3390/w13243542, 2021.

Constanza, R. and Ruth, M.: Using Dynamic Modeling to Scope Environmental Problems and Build Consensus, Environ. Manage., 22, 183–195, https://doi.org/10.1007/s002679900095, 1998.

de Oliveira Ferreira Silva, C.: The Challenge of Model Validation and Its (Hydrogeo) ethical Implications for Water Security, In Computational Intelligence for Water and Environmental Sciences, pp. 477-489, Singapore: Springer Nature Singapore, 2022.

Deitrick, A. R., Torhan, S. A., and Grady, C. A.: Investigating the Influence of Ethical and Epistemic Values on Decisions in the Watershed Modeling Process, Water Resour. Res., 57, https://doi.org/10.1029/2021WR030481, 2021.

Demeritt, D. The Construction of Global Warming and the Politics of Science. Ann Am Assoc Geogr, 91(2), 307–337, 2001.

Dobson, B., Wagener, T., and Pianosi, F.: How Important Are Model Structural and Contextual Uncertainties when Estimating the Optimized Performance of Water Resource Systems?, Water Resour. Res., 55, 2170–2193, https://doi.org/10.1029/2018WR024249, 2019.

Doorn, N.: Responsibility Ascriptions in Technology Development and Engineering: Three Perspectives, Sci. Eng. Ethics, 18, 69–90, https://doi.org/10.1007/s11948-009-9189-3, 2012.

Étienne, M.: Companion modelling: a participatory approach to support sustainable development. Editions Quae, 368 p., 2011

Falconi, S. M. and Palmer, R. N.: An interdisciplinary framework for participatory modeling design and evaluation—What makes models effective participatory decision tools?, Water Resour. Res., 53, 1625–1645, https://doi.org/10.1002/2016WR019373, 2017.

Fernandez, S.: Much Ado About Minimum Flows…Unpacking indicators to reveal water politics, Geoforum, 57, 258–271, https://doi.org/10.1016/j.geoforum.2013.04.017, 2014.

Garcia-Cuerva, L., Berglund, E. Z., and Rivers, L.: Exploring Strategies for LID Implementation in Marginalized Communities and Urbanizing Watersheds, in: World Environmental and Water Resources Congress 2016, World Environmental and Water Resources Congress 2016, West Palm Beach, Florida, 41–50, https://doi.org/10.1061/9780784479889.005, 2016.

Giglioli, I. and Swyngedouw, E.: Let's Drink to the Great Thirst! Water and the Politics of Fractured Techno-natures in Sicily: Water and the politics of fractured techno-natures in Sicily, Int. J. Urban Reg. Res., 32, 392–414, https://doi.org/10.1111/j.1468-2427.2008.00789.x, 2008.

Godinez-Madrigal, J., Van Cauwenbergh, N., and van der Zaag, P.: Production of competing water knowledge in the face of water crises: Revisiting the IWRM success story of the Lerma-Chapala Basin, Mexico, Geoforum, 103, 3–15, https://doi.org/10.1016/j.geoforum.2019.02.002, 2019.

Golinski J.:Making Natural Knowledge: Constructivism and the History of Science. Chicago, IL: University of Chicago Press, 2005.

Haas, P. M.: Introduction: Epistemic Communities and International Policy Coordination, 36, 1992.

Haeffner, M., Jackson-Smith, D., and Flint, C. G.: Social Position Influencing the Water Perception Gap Between Local Leaders and Constituents in a Socio-Hydrological System, Water Resour. Res., 54, 663–679, https://doi.org/10.1002/2017WR021456, 2018.

Haeffner, M., Hellman, D., Cantor, A., Ajibade, I., Oyanedel-Craver, V., Kelly, M., Schifman, L., and Weasel, L.: Representation justice as a research agenda for socio-hydrology and water governance, Hydrol. Sci. J., 66, 1611–1624, https://doi.org/10.1080/02626667.2021.1945609, 2021.

Hamilton, S.H., Pollino, C.A., Stratford, D.S., Fu, B. and Jakeman, A.J.: Fit-for-purpose environmental modeling: Targeting the intersection of usability, reliability and feasibility. Environmental Modelling & Software, 148, p.105278, 2022.

Haines, S.: Reckoning Resources: Political Lives of Anticipation in Belize's Water Sector, Technol. Stud., 32, 22, 2019.

Haraway, D.: Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective, Fem. Stud., 14, 575, https://doi.org/10.2307/3178066, 1988.

Haraway, D. J.: Manifesto for cyborgs: science, technology, and socialist feminism in the 1980s, Social. Rev., 80, 65–108, 1985.

Harmel, R.D., Smith, P.K., Migliaccio, K.W., Chaubey, I., Douglas-Mankin, K.R., Benham, B., Shukla, S., Muñoz-Carpena, R. and Robson, B.J.: Evaluating, interpreting, and communicating performance of hydrologic/water quality models considering intended use: A review and recommendations. Environ Model Softw., 57, 40-51, 2014.

Harvey, F. and Chrisman, N.: Boundary Objects and the Social Construction of GIS Technology, Environ. Plan. Econ. Space, 30, 1683–1694, https://doi.org/10.1068/a301683, 1998.

Hasala, D., Supak, S., and Rivers, L.: Green infrastructure site selection in the Walnut Creek wetland community: A case study from southeast Raleigh, North Carolina, Landsc. Urban Plan., 196, 103743, https://doi.org/10.1016/j.landurbplan.2020.103743, 2020.

Holländer, H. M., Bormann, H., Blume, T., Buytaert, W., Chirico, G. B., Exbrayat, J.-F., Gustafsson, D., Hölzel, H., Krauße, T., Kraft, P., Stoll, S., Blöschl, G., and Flühler, H.: Impact of modellers' decisions on hydrological a priori predictions, Hydrol. Earth Syst. Sci., 18, 2065–2085, https://doi.org/10.5194/hess-18-2065-2014, 2014.

Holifield, R.: How to speak for aquifers and people at the same time: Environmental justice and counter-network formation at a hazardous waste site. Geofor., 40(3), 363-372, 2009.

Holifield R.: Environmental justice as recognition and participation in risk assessment: negotiating and translating health risk at a superfund site in Indian country. Ann Assoc Am Geographers. 102:591–613, 2012.

Hulme, Mike.: Why we disagree about climate change: Understanding controversy, inaction and opportunity. Cambridge University Press, Cambridge, 432 pp., https://doi.org/10.1017/CBO9780511841200, 2009.

Jackson, S.: Water models and water politics: design, deliberation, and virtual accountability, Proc Int Conf Digit Gov Res, 95-104, 2006.

Jasanoff, S. and Kim, S.-H.: Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea, Minerva, 47, 119–146, https://doi.org/10.1007/s11024-009-9124-4, 2009.

Jenkins, D. G. and McCauley, L. A.: GIS, SINKS, FILL, and Disappearing Wetlands: Unintended Consequences in Algorithm Development and Use, SAC'06, Dijon, France, 6, 2006.

Jensen, C. B.: A flood of models: Mekong ecologies of comparison, Soc. Stud. Sci., 50, 76–93, https://doi.org/10.1177/0306312719871616, 2020.

Junier, S. J.: Modelling expertise: Experts and expertise in the implementation of the Water Framework Directive in the Netherlands, Delft University of Technology, https://doi.org/10.4233/UUID:EEA8A911-F786-4158-A67E-B99663275BF8, 2017.

King, J. L. and Kraemer, K. L.: Models, Facts, and the Policy Process: The Political Ecology of Estimated Truth, Center for Research on Information Systems and Organizations, Irvine, 1993.

Knorr-Cetina, K.: Epistemic cultures: how the sciences make knowledge, Harvard University Press, Cambridge, Mass, 329 pp., 1999.

Kouw, M.: Standing on the Shoulders of Giants—And Then Looking the Other Way? Epistemic Opacity, Immersion, and Modeling in Hydraulic Engineering, Perspect. Sci., 24, 206–227, https://doi.org/10.1162/POSC_a_00201, 2016.

Kouw, M.: Risks in the Making: The Mediating Role of Models in Water Management and Civil Engineering in the Netherlands, Berichte Zur Wiss., 40, 160–174, https://doi.org/10.1002/bewi.201701823, 2017.

Kroepsch, A. C.: Groundwater Modeling and Governance: Contesting and Building (Sub)Surface Worlds in Colorado's Northern San Juan Basin, Engag. Sci. Technol. Soc., 4, 43–66, https://doi.org/10.17351/ests2018.208, 2018.

Krueger, T. and Alba, R.: Ontological and epistemological commitments in interdisciplinary water research: Uncertainty as an entry point for reflexion. Frontiers in Water, 4, p.1038322, 2022.

Krueger, T., Maynard, C., Carr, G., Bruns, A., Mueller, E. N., and Lane, S.: A transdisciplinary account of water research, WIREs Water, 3, 369–389, https://doi.org/10.1002/wat2.1132, 2016.

Laborde, S.: Environmental Research from Here and There : Numerical Modelling Labs as Heterotopias, Environ. Plan. Soc. Space, 33, 265–280, https://doi.org/10.1068/d14128p, 2015.

Ländstrom, C., Whatmore, S. J., and Lane, S. N.: Virtual engineering: computer simulation modelling for flood risk management in England, Sci. Technol. Stud., 24, 3–22, 2011a.

Landström, C., Whatmore, S. J., Lane, S. N., Odoni, N., Ward, N., and Bradley, S.: Coproducing flood risk knowledge: redistributing expertise in critical "participatory modelling," Environ. Plan. Econ. Space, 43, 1616–1633, 2011b.

Lane, S.N., Odoni, N., Landström, C., Whatmore, S.J., Ward, N. and Bradley, S.: Doing flood risk science differently: an experiment in radical scientific method. Transactions of the Institute of British Geographers, 36(1),15-36, 2011a.

Lane, S. N., Landström, C., and Whatmore, S. J.: Imagining flood futures: risk assessment and management in practice, Philos. Trans. R. Soc. Math. Phys. Eng. Sci., 369, 1784–1806, https ://doi.org/10.1098/rsta.2010.0346, 2011b.

Lane, S.N., 2012. Making mathematical models perform in geographical space(s). Chapter 17 in Agnew, J. and Livingstone, D. Handbook of Geographical Knowledge. Sage, London

Lane, S.N., November, V., Landström, C. and Whatmore, S.: Explaining rapid transitions in the practice of flood risk management. Annals of the Association of American Geographers, 103(2), 330-342, 2013.

Lane, S. N.: Acting, predicting and intervening in a socio-hydrological world, Hydrol. Earth Syst. Sci., 18, 927–952, https://doi.org/10.5194/hess-18-927-2014, 2014.

Lasswell, Harold. D.: Politics: Who Gets What, When, How, Whittlesey House, Cleveland, New York, 264, https://doi.org/10.1007/978-3-531-90400-9_60, 1936.

Latour, B.: When things strike back: a possible contribution of "science studies" to the social sciences, Br. J. Sociol., 51, 107–123, 2000.

Latour, B.: The promises of constructivism, in Ihde D. (Ed) Chasing Technology : Matrix of Materiality, Indiana Series for the Philosophy of Science, Indiana University Press, 27-46, 2003.

Latour, B. and Woolgar, S.: Laboratory life: the construction of scientific facts, Princeton University Press, Princeton, N.J, 294 pp., 1986.

Law, J., After method: Mess in social science research. Psychology Press, 2004.

Leigh Star, S.: This is Not a Boundary Object: Reflections on the Origin of a Concept, Sci. Technol. Hum. Values, 35, 601–617, https://doi.org/10.1177/0162243910377624, 2010.

Linton, J.: Modern water and its discontents: a history of hydrosocial renewal, WIREs Water, 1, 111–120, https://doi.org/10.1002/wat2.1009, 2014.

Linton, J. and Budds, J.: The hydrosocial cycle: Defining and mobilizing a relational-dialectical approach to water. Geoforum, 57, 170-180, 2014.

Litfin, K. T.: The Gendered Eye in the Sky: A Feminist Perspective on Earth Observation Satellites, Front. J. Women Stud., 18, 26, https://doi.org/10.2307/3346964, 1997.

Losee, R. M.: A discipline independent definition of information, J. Am. Soc. Inf. Sci., 48, 254–269, https://doi.org/10.1002/(SICI)1097-4571(199703)48:3<254::AID-ASI6>3.0.CO;2-W, 1997.

MacKenzie, D. A. and Wajcman, J. (Eds.): The social shaping of technology, 2nd ed., Open University Press, Buckingham [Eng.] ; Philadelphia, 462 pp., 1999.

MacKenzie D.: An engine, not a camera: How financial models shape markets. The MIT Press, Cambridge, 392 pp., 2008.

Macnaghten, P.: Governing Science and Technology: From the Linear Model to Responsible Research and Innovation, in: The Cambridge Handbook of Environmental Sociology, edited by: Legun, K., Keller, J., Bell, M., and Carolan, M., Cambridge University Press, 347–361, https://doi.org/10.1017/9781108554510.023, 2020.

Maeda, E. E., Haapasaari, P., Helle, I., Lehikoinen, A., Voinov, A., and Kuikka, S.: Black Boxes and the Role of Modeling in Environmental Policy Making, Front. Environ. Sci., 9, 629336, https://doi.org/10.3389/fenvs.2021.629336, 2021.

Meenar, M., Fromuth, R. and Soro, M.: Planning for watershed-wide flood-mitigation and stormwater management using an environmental justice framework. Environmental Practice, 20(2-3), 55-67, 2018.

Melsen, L. A.: It Takes a Village to Run a Model—The Social Practices of Hydrological Modeling, Water Resour. Res., 58, https://doi.org/10.1029/2021WR030600, 2022.

Melsen, L. A., Addor, N., Mizukami, N., Newman, A. J., Torfs, P. J. J. F., Clark, M. P., Uijlenhoet, R., and Teuling, A. J.: Mapping (dis)agreement in hydrologic projections, Hydrol. Earth Syst. Sci., 22, 1775–1791, https://doi.org/10.5194/hess-22-1775-2018, 2018a.

Melsen, L. A., Vos, J., and Boelens, R.: What is the role of the model in socio-hydrology? Discussion of "Prediction in a socio-hydrological world," Hydrol. Sci. J., 63, 1435–1443, https://doi.org/10.1080/02626667.2018.1499025, 2018b.

Melsen, L. A., Teuling, A. J., Torfs, P. J. J. F., Zappa, M., Mizukami, N., Mendoza, P. A., Clark, M. P., and Uijlenhoet, R.: Subjective modeling decisions can significantly impact the simulation of flood and drought events, J. Hydrol., 568, 1093–1104, https://doi.org/10.1016/j.jhydrol.2018.11.046, 2019.

Mendoza, P. A., Clark, M. P., Mizukami, N., Gutmann, E. D., Arnold, J. R., Brekke, L. D., and Rajagopalan, B.: How do hydrologic modeling decisions affect the portrayal of climate change impacts?, Hydrol. Process., 30, 1071–1095, https://doi.org/10.1002/hyp.10684, 2016.

Molle, F.: Nirvana Concepts, Narratives and Policy Models: Insights from the Water Sector, 1, 26, 2008.

Morgan, M. S., & Morrison, M. (Eds.): Models as mediators: Perspectives on natural and social science (No. 52). Cambridge University Press, Cambridge. 401 pp., 1999.

Munk, A. K.: Risking the Flood: Cartographies of Things to Come, University of Oxford, Oxford, 268 pp., 2010.

Nearing, G.S., Kratzert, F., Sampson, A.K., Pelissier, C.S., Klotz, D., Frame, J.M., Prieto, C. and Gupta, H.V.: What role does hydrological science play in the age of machine learning?. Water Resources Research, 57(3), p.e2020WR028091, https://doi.org/10.1029/2020WR028091, 2021.

Opitz-Stapleton, S. and MacClune, K.: Chapter 11 Scientific and Social Uncertainties in Climate Change: The Hindu Kush-Himalaya in Regional Perspective, in: Community, Environment and Disaster Risk Management, vol. 11, edited by: Lamadrid, A. and Kelman, I., Emerald Group Publishing Limited, 207–237, https://doi.org/10.1108/S2040-7262(2012)0000011017, 2012.

Packett, E., Grigg, N. J., Wu, J., Cuddy, S. M., Wallbrink, P. J., and Jakeman, A. J.: Mainstreaming gender into water management modelling processes, Environ. Model. Softw., 127, 104683, https://doi.org/10.1016/j.envsoft.2020.104683, 2020.

Pielke Jr. R.A.: The honest broker: making sense of science in policy and politics. Cambridge University Press, New York, 188 pp., https://doi.org/10.1017/CBO9780511818110, 2007.

Puy, A., Sheikholeslami, R., Gupta, H. V., Hall, J. W., Lankford, B., Lo Piano, S., Meier, J., Pappenberger, F., Porporato, A., Vico, G., and Saltelli, A.: The delusive accuracy of global irrigation water withdrawal estimates, Nat. Commun., 13, 3183, https://doi.org/10.1038/s41467-022-30731-8, 2022.

Rainwater, K., Stovall, J., Frailey, S., and Urban, L.: Transboundary Impacts on Regional Ground Water Modeling in Texas, Ground Water, 43, 706–716, https://doi.org/10.1111/j.1745-6584.2005.00068.x, 2005.

Ramsey, K.: GIS, modeling, and politics: On the tensions of collaborative decision support, J. Environ. Manage., 90, 1972–1980, https://doi.org/10.1016/j.jenvman.2007.08.029, 2009.

Rusca, M. and Di Baldassarre, G.: Interdisciplinary Critical Geographies of Water: Capturing the Mutual Shaping of Society and Hydrological Flows, Water, 11, 1973, https://doi.org/10.3390/w11101973, 2019.

Saltelli, A., Bammer, G., Bruno, I., Charters, E., Di Fiore, M., Didier, E., Nelson Espeland, W., Kay, J., Lo Piano, S., Mayo, D. and Pielke Jr, R.: Five ways to ensure that models serve society: a manifesto, Nature, 2020.

Saltelli, Andrea, and Monica Di Fiore, eds. The Politics of Modelling: Numbers Between Science and Policy. Oxford University Press, pp. 231, 2023.

Sanz, D., Vos, J., Rambags, F., Hoogesteger, J., Cassiraga, E., and Gómez-Alday, J. J.: The social construction and consequences of groundwater modelling: insight from the Mancha Oriental aquifer, Spain, Int. J. Water Resour. Dev., 35, 808–829, https://doi.org/10.1080/07900627.2018.1495619, 2019.

Shrader-Frechette, K.: Hydrogeology and framing questions having policy consequences, Philosophy of Science, 64(S4), S149-S160, 1997.

Sismondo, S.: Models, Simulations, and Their Objects, Sci. Context, 12, 247–260, https://doi.org/10.1017/S0269889700003409, 1999.

Srinivasan, V., Sanderson, M., Garcia, M., Konar, M., Blöschl, G., and Sivapalan, M.: Prediction in a socio-hydrological world, Hydrol. Sci. J., 1–8, https://doi.org/10.1080/02626667.2016.1253844, 2016.

Srinivasan, V., Sanderson, M., Garcia, M., Konar, M., Blöschl, G., and Sivapalan, M.: Moving socio-hydrologic modelling forward: unpacking hidden assumptions, values and model structure by engaging with stakeholders: reply to "What is the role of the model in socio-hydrology?," Hydrol. Sci. J., 63, 1444–1446, https://doi.org/10.1080/02626667.2018.1499026, 2018.

Star, S. L. and Griesemer, J. R.: Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39, Soc. Stud. Sci., 19, 387–420, https://doi.org/10.1177/030631289019003001, 1989.

Stengers, I., Another science is possible: A manifesto for slow science. John Wiley & Sons, 2018.

Thompson, E. L. and Smith, L. A.: Escape from model-land, Economics, 13, 20190040, https://doi.org/10.5018/economics-ejournal.ja.2019-40, 2019.

Trombley, J. M.: An Environmental Anthropology of Modeling and Management on the Chesapeake Bay Watershed, University of Maryland, College Park, 2017.

Turnhout, E., Hisschemöller, M., and Eijsackers, H.: Ecological indicators: between the two fires of science and policy. Ecol. Indic., 7(2), 215-228. https://doi.org/10.1016/j.ecolind.2005.12.003, 2007

Turner, M. D. Production of environmental knowledge: Scientists, complex natures, and the question of agency. Knowing Nature: Conversations at the Intersection of Political Ecology and Science Studies. University of Chicago Press, Chicago, 25–29, 2011.

Venot, J.-P., Vos, J., Molle, F., Zwarteveen, M., Veldwisch, G. J., Kuper, M., Mdee, A., Ertsen, M., Boelens, R., Cleaver, F., Lankford, B., Swatuk, L., Linton, J., Harris, L. M., Kemerink-Seyoum, J., Kooy, M., and Schwartz, K.: A bridge over troubled waters, Nat. Sustain., https://doi.org/10.1038/s41893-021-00835-y, 2021.

Voinov, A., Kolagani, N., McCall, M. K., Glynn, P.D., Kragt, M. E., Ostermann, F. O., Pierce, S. A., Ramu, P.: Modelling with stakeholders – Next generation. Environ Model Softw, 77, 196-220. https://doi.org/10.1016/j.envsoft.2015.11.016, 2016.

Voinov, A., Seppelt, R., Reis, S., Nabel, J.E. and Shokravi, S.:Values in socio-environmental modelling: Persuasion for action or excuse for inaction. Environ Model Softw, 53, 207-212, http://dx.doi.org/10.1016/j.envsoft.2013.12.005, 2014.

Wesselink, A., de Vriend, H., Barneveld, H., Krol, M. and Bijker, W.: Hydrology and hydraulics expertise in participatory processes for climate change adaptation in the Dutch Meuse. Water Sci Technol, 60(3), 583-595, 2009.

Wesselink, A., Kooy, M., and Warner, J.: Socio-hydrology and hydrosocial analysis: toward dialogues across disciplines. WIREs Water, 4, https://doi.org/10.1002/wat2.1196, 2017.

Whatmore, S.J. and Landström, C.: Manning's N: Putting roughness to work. In Howlett, P. and Morgan, M.S. eds., How well do facts travel?: The dissemination of reliable knowledge. Cambridge University Press, 111-135, 2010.

Wheeler, K. G., Hall, J. W., Abdo, G. M., Dadson, S. J., Kasprzyk, J. R., Smith, R., and Zagona, E. A.: Exploring Cooperative Transboundary River Management Strategies for the Eastern Nile Basin, Water Resour. Res., 54, 9224–9254, https://doi.org/10.1029/2017WR022149, 2018.

Wheeler, K.G., Robinson, C.J. and Bark, R.H.: Modelling to bridge many boundaries: the Colorado and Murray-Darling River basins. Reg Environ Change, 18, 1607-1619, 2018.

Woolgar, S. and Cooper, G., Do artefacts have ambivalence: Moses' bridges, winner's bridges and other urban legends in S&TS. Social studies of science, 29(3), 433-449, 1999.

Zwarteveen, M., Kemerink-Seyoum, J. S., Kooy, M., Evers, J., Guerrero, T. A., Batubara, B., Biza, A., Boakye-Ansah, A., Faber, S., Cabrera Flamini, A., Cuadrado-Quesada, G., Fantini, E., Gupta, J., Hasan, S., ter Horst, R., Jamali, H., Jaspers, F., Obani, P., Schwartz, K., Shubber, Z., Smit, H., Torio, P., Tutusaus, M., and Wesselink, A.: Engaging with the politics of water governance, WIREs Water, 4, https://doi.org/10.1002/wat2.1245, 2017.

In this letter we summarise how we have revised the article based on the feedback of the reviewers and community. Elaborate responses to each point raised by the reviewers are attached as annex.

1. **Introduction**
- We have included more explicit references to already ongoing debates on water models, in response to reviewer #1 and #3, including the fit-for purpose discussion.
- We have added a discussion on why it is necessary to discuss the influence of water models in particular, especially in relation to their connection with water infrastructure and the role of water in society, as well as the few articles that exist on the topic.

2. **Defining models**
- We have elaborated on work done by Science-Technology and Society scholars specifically to understand the influence of technology and knowledge, including the work of Demeritt and Stengers in response to reviewer #2
- We have made clearer in the revised article that the political nature of modelling is always an interplay of actors (commissioners, scientists, modelers, politicians, other stakeholders) and the model itself, in response to reviewer #1 and #3.

3. **Methodology**
- In response to all reviewers, we have clarified why we find the combination of a narrative review and the ROSES method for review useful in this case, despite the large number of articles that have been excluded.
- In response to reviewer #3, we have more clearly explained why certain articles are included and others excluded.
- We have adjusted the query to include hydraulic and groundwater models too
- We have reviewing the articles suggested by the three reviewers and community, to identify how/if they can be included in the review. We are thankful for the additional input.
- We have clarified how the articles were analysed and how the four overarching themes and 13 topics were found.

**4. Results**
- We have updated the results based on the additional articles that will be included in the review.

**5. & 7. Discussion and conclusion**
- We have streamlined the discussion to avoid repetition between the conclusion and discussion, based on feedback for reviewers #2 and #3.

**6. Call for power-sensitive modelling**
- A very important element of our revision is to add examples to clarify and exemplify the considerations that can guide power-sensitive modelling, in response to reviewers #2 and #3, especially referring to existing work.

Point-by-point response to the reviews including a list of all relevant changes made in the manuscript

**Reviewer 1**

**Reviewer:** *Because the authors seem to be falling into the trap of talking about models as if they were responsible for outcomes – and as I have written several times before, decision makers are only too happy to transfer responsibility to the models. But that is not how (in general) it works. With the exception of some academic studies (such as many of those cited here) models are conditioned for a purpose. The power relationship here is not at all in the model, the model is only a consequence of making assumptions, so the important factor is who decides on the assumptions. This will be constrained initially by how a project is framed (the political as used by the authors, which may well reflect the vested interests of whoever is doing the commissioning or decision making) and then by how the modeller interprets the brief (including how much time is costed when bids are competitive). Of course, in a sociohydrological model, then the relative importance of different stakeholder groups might be a factor that needs to be included – but that again comes down to whether the assumptions made are appropriate for the purpose.*

**Answer:** We thank Keith Beven for the review, and for pointing out many commonalities in understanding the political processes of developing and using models while challenging our argument that the model itself is not neutral. We find this challenge valuable because it is a common point of alienation between hydrologists and critical water scholars that we wish to constructively engage with in this article.

We here summarise how we will deepen our argument in the revised article:
We first emphasise that the non-neutrality of models and their power is not necessarily a negative trait, and that a model is not all-powerful. Its power manifests in different ways, some very visible and some not, and sometimes successful and sometimes less. For instance, models can legitimise and/or challenge policy projects; they provide certain pathways for action and potentially exclude others, and they reify assumptions about the world. Even if a model is only used in academics it still informs the way we think about those water-related issues. Models can make it challenging to question certain assumptions in a model, especially as choices are encoded in the model and become "black boxed", which makes it increasingly difficult to unravel, especially for non-experts. They can also work in an in- or excluding way, including through certain jargon and language, or certain technology and visualizations of the results used.

Second point is that models are not just a partial, but also a specific representation of reality. People involved in the modelling process might have different ideas of what variables, boundaries, scales etc. are relevant or not (depending on their knowledge, values and experiences) in light of the defined purpose of the model. So, whether a model is "fit for purpose" will be assessed differently by different actors. Modelling involves many (conscious and unconscious) choices and assumptions, which are the product of – and in turn influence -the interplay between different actors (commissioners, users, modelers and affected stakeholders). Each bring in different expertise, world views and ideas for the future. This interplay is enabled or constrained by technology and the modelling process.

Third, not all stakeholders have the same ability to influence the modelling choices or use the modelling outcomes. It is therefore that we wish to reconnect both model developers and users with the models so that they recognize the possibilities that the model enables and those that it forecloses and to engage with this constructively.

In the attached file we share our answer to the other points made by the reviewer.

**Reviewer:** *The confounding of this distinction starts in the section on Defining Models which suggests that a model is anything that takes an input to produce an output. There is nothing here on purpose, fitness-for-purpose, responsibility or explicitly defining the condition tree of assumptions (including*

*considerations of uncertainty) in any model application. Indeed, the study cited by Ramsey (2009) (L165ff) is an example of where the assumptions used made the model not fit-for-purpose (see the discussion of fit-for-purpose in Beven and Lane, 2022). The later example of Andersson et al (2004) is also an example of an application where the model was not fit for purpose. So is not part of the issue here the question of how to get consensus on what might be fit-for-purpose (depending on the purpose). It could be useful here to bring in the idea of the condition tree of assumptions as a way of improving this process – both to get participatory agreement and understanding (and an audit trail for later reconsideration). See for example, Beven et al. CIRIA 2014; or the Page et al., HESS 2023 toolbox). This is a way of getting the more open/explicit approach suggested in Section 4.1.3 and in the conclusions.*

**Answer:** We agree that purpose and fitness-for-purpose are important considerations in the practice of building and using models that we will discuss more explicitly in the revised article. The contestations about the purpose of a model and how that translates into the model choices, and the fitness-for-purpose of a model show the political nature of the model and the modelling process itself. We like the idea of an audit trail, to which we may add our aforementioned political considerations. For instance, what a power-sensitive modelling approach could contribute to the principles you share 'For a Turing-liketest for model plausibility/(in)validation and fitness-for-purpose' in Beven and Lane (2022: pp 7) is that it could help to guide additional questions such as 'whose purpose' is served, and for what? (related to point 1). And what is relevant expertise to take into account, and does this only concern environmental modellers, or also other types of knowledge? (related to point 7).

**Reviewer:** *I would argue that the political in the sense used by the authors is not really the "broad influence of the models" that they suggest but rather the whole framing of the decision making process, within which the model is only a tool. This can often be seen in how historical legal constraints on water resources management dominate any attempts to achieve sustainability of use (for either human use or biodiversity) regardless of whatever model might be used. I accept, of course, that models are not necessarily neutral in this process, even if most modellers that I know of will try to do the best they can given the data they might have available). The assumptions can certainly reflect the power of requiring a certain outcome (a nice example, in the inquiry into the Sellafield Rock Experiment Facility as the first stage in the UK nuclear waste disposal strategy, where the two opposing sides used quite different assumptions in assessing groundwater flow pathways).*

**Answer:** We agree that models are connected to a wider political process of managing water. But models are not just tools because they encode a certain mindset of how to understand and manage water as described above. In fact, the very notion that they are just tools, which is very prominent in water management, obscures the many choices in structuring understanding and management of water (e.g. the optimisation paradigm further discussed below) that are inscribed in the models. In that sense they are even more politically powerful in pushing a particular discourse because they are portrayed as being neutral. The Sellafield example is a good one, although we are also interested in the more nuanced (everyday) workings of model politics.

**Reviewer:** *The idea of purpose also interacts with framing – this is evident in the examples of the Rhone and Seine in Section 4.1.2. These projects were tasked with different purposes when the projects were commissioned – it is too simple to say that this was the result of contrasting world-views of geographers and engineers. It was actually the result of upstream decisions about commissioning the projects before any model or modeller was even involved (actually that might not be entirely correct, as I believe there was initially a commission project at least on the Seine to consider what might be technologically possible - but more background detail would be needed to make that point). It is again the wider political framework that is determining outcomes here – not the model (though, yes, as above for the Seine, in some*

*cases the available model technology might feedback to how projects are defined, but that is not what you are saying here. It would be an interesting study in itself – many of the commercial SHE model projects were of that type, for example).*

**Answer:** We shall make clearer in the revised article that the political nature of modelling is always an interplay of actors (commissioners, scientists, modelers, politicians, other stakeholders) and the model itself. Bouleau (2014) describes in the article on the Seine and the Rhone how management-cultures shape how models are designed, and in turn how models shape their environments. It is thus both the (also historical) wider political framework that is described in this article, and how it manifests in and through the model, as well as the influence from the models themselves. Especially the articles reviewed under section 4.3 'Modelling and real-world impact' showcases this interaction.

**Reviewer:** *L265 – you put 'optimal' in commas but without further comment when there has long been discussion in hydrological modelling critical of the concept of optimality, either in model outputs or in decision making (e.g. my 2006 equifinality paper and earlier). So yes, modeller choices matter and make a difference but these are old examples now – better if multiple model outputs are considered (model ensembles are now widely used) in a context of assessing fitness-for-purpose within the limitations of uncertainty.*

**Answer:** We shall explain the commas in the article. What we add is that even with ensemble models the political charge of models does not go away – we will take up this point in an expanded discussion. Although these are old discussions, and perhaps an old example, the concept of optimality is still alive and is relevant to discuss in the context of the review. Also, few scientific articles describe the impact of this modelling practice in a case study. More openness on the effects of these modelling choices could contribute to changing these practices.

We also fully agree that it is indeed better when multiple sources of knowledge, including models, are used for decision making. This also hints at the different influences different models can have.

**Reviewer:** *L297. – yes, but there was a history to that in a simplified interpretation of the results of a CEH study of the impacts of climate change on the Severn and Thames catchments. It was conditional on the climate scenarios chosen and available at that time (and so should have been subject to revision). It is not what the models or modellers of that study said but, for simplicity, (and cheapness in low budget applications) it has persisted. Another example of the dominance of the wider political framework over the model.*

**Answer:** We are grateful for these further insights. Still, whatever wider political economy and modellers' choices played a role, they are now encoded in that particular model and not so easy to unpack (hence your insider knowledge is so valuable).

**Reviewer:** *Section 4.3.1. OK, so models can be used in ways to support vested interests but that again is a political issue, not a model-specific issue (as in the Sellafield example above where the opposing side effectively won the case by demonstrating the huge potential uncertainty in model results. Similar model, different assumptions. So why is the comment about models being value neutral just thrown in at the end without further discussion? That is a claim made for political ends, not a feature of the model.*

**Answer:** This too we shall make clearer in the revised article. Models have characteristics that make them have specific and differentiating effects for different stakeholders (through the model outcomes), and allow different people to participate in the modelling process, use the model and model outcomes differently.

If models are purported as being value neutral when in fact they are not then the models are implicated in this political maneuvering.

*Reviewer: L335. In respect of groundwater models there are also examples of studies showing that different experts come up with different conceptual models (e.g. Refsgaard conceptual groundwater model paper), but, more important here, are the post-audit studies of Bredehoeft and Konikow (1992) and Anderson and Woesnner (1992) in a special issue of AWR who showed that nearly all groundwater models did not prove to be correct with hindsight, but mostly because the projected boundary conditions had not proven correct (ie. the assumptions rather than the model again). That has put people off doing post-audit analyses for both hydrological and groundwater models ever since….. (even though that could be an invaluable learning process).*

**Answer:** Thank you very much for bringing this up. We agree with the importance of post-audit of models, and doing that together with affected stakeholders. We would see the assumptions as part of the model, not something outside of the model.

As we state in line 335: "Kroepsch (2018) and Sanz et al. (2019) discussed how groundwater models can be used to legitimise policies even if there is limited information available." It is telling that there are few post audit analyses done for these kind of projects indeed, and a point of attention for our joint project to improve models. We would call for a power-sensitive post-audit, both to learn how to improve the representation of the physical – including the boundary conditions as you point out, as well as to reflect on the influence the model had and whether this was constructive in the decision making process. We will add a reflection to this end in the discussion (section 4.5).

*Reviewer: Section 4.3.2. "When modelling is presented as a neutral scientific process …. legitimacy given by external consultants." These are surely not the same thing as is being implied here. The consultants may well have been trying to do the best job possible given the data available, independently of any vested interest (you do not provide any evidence to the contrary). The model, given better data might well have been a better representation of the system. The model can be neutral in that respect; even if the way it is used might not be. The question again is whether that implementation should have been considered fit-for-purpose for the decisions being made. "Framing their actions as illegal…" also surely has nothing to do with the model?*

**Answer:** To us it's again about the model's influence in these situations. Other comments made above apply.

*Reviewer: L363 "The decision over water allocation was eventually enforced through influence at the highest political level, the President of Mexico. Jensen (2020) also confirmed that the power of high-level decision makers plays a key role. In the case of the Mekong, the author showed there is a certain saturation in knowledge developed by models, and there is a clear limitation in their impact as governments were unwilling to build on these insights." But exactly! That is not the same as you then go on to say "The previous examples show how models can work exclusively…" when their outputs are rejected or ignored????? The models are not working exclusively in support of the chosen solution surely? And what you discuss in the remainder of this section is how a model might be useful as a tool in framing good practice (as was also the case of Pickering). But it is again not the model as such (which might still be too much of a simplification or lacking in data), but the way the model is used.*

**Answer:** We will rephase this section to strengthen the consistency of the argument. Indeed, the influence of models is relative: models will not automatically produce outcomes. However, models and model outcomes do make certain usages possible and disable others.

*Reviewer: Section 4.4.1 You mention the importance of scale here but not the importance of visualisations in how local people can interact with model outputs (as in the models of everywhere concepts of Beven, HESS 2007, Beven et al., JRBM 2014; and Blair et al., EMS 2019).*

**Answer:** That is indeed an omission in the article, and thank you for pointing it out. How the output of a model has influence, for instance through visualisations, graphs, scales or perhaps interactive dashboards, has not specifically come up in the articles we reviewed. It is interesting indeed, as there are ample articles that point to the politics of maps, or for instance the influence of a certain use of colour.

*Reviewer: Avoid models that are overly complex? Over-complex with respect to what? To the problem at hand or to the understanding of stakeholders. The second should surely not override the adequate complexity of the first?*

**Answer:** The advice of Opitz-Stapleton and MacClune (2012) to reduce model complexity is in relation to connecting and discussing the model with stakeholders. There are cases in which researchers have chosen to work with simpler models to enable stakeholders to engage in the analysis, for instance also suggested by Srivinasan et al. (2018) of which the article is included in the review. In the revised version we will clarify this point.

*Reviewer: Having said that, the power relationships of this type of co-evolutionary modelling are definitely an issue, but as some of your example studies have shown not necessarily insurmountable if the will is there. But many of the problems you have identified so far are a result of the imposition of power structures, regardless of the model or its results.*

**Answer:** Comments above apply. Those power structures are inscribed in the models.

*Reviewer: Section 4.4.2. Consideration of assumptions about uncertainty (including epistemic uncertainties) in modelling is critical to model evaluation and fitness-for-purpose but this is only really mentioned in this Section in relation to an application of SWAT (widely applied – it is free to use - but which elsewhere has been shown to be not fit-for-purpose for the type of application described here, even allowing for uncertainty – see Hollaway et al., JH, 2018).*

**Answer:** The main point we want to show is that the characteristics of the model influence the interaction between the model, model users and affected stakeholders. Each stakeholder will make a different assessment of the fitness-for-purpose of the model and the treatment of the uncertainties. That process makes the model political.

*Reviewer: What do the authors mean here by "As modelling inhibits more uncertainty than measurement" (L431)? The model here is being used to predict years ahead so should surely be more uncertain than any available measurements (that cannot in any case be made in the future).*

**Answer:** The sentence refers to the summary of Wardropper et al. (2017) on how inherent uncertainty in the Soil and Water Assessment Tool (SWAT) application to the Yahara Watershed in Wisconsin (USA) influenced the development and implementation of a water quality management program. It is indeed a redundant sentence, which we have adjusted into: "The authors questioned how the inherent uncertainty in this approach affected people in the watershed."

*Reviewer: L447 Pickering – "ultimately played a key role in shaping flood management strategy in the area." Well yes – but you should then perhaps finish the story. It showed that the NFM strategy preferred by local stakeholders would not protect the properties at risk. The fact that the EA had also been involved in the process then meant that they invested £1.5m in a concrete flood detention basin, despite the cost-*

*benefit for the scheme being considered too low compared with other sites. That was then a political decision.*

**Answer:** We appreciate these further insights. Again, different models were implicated here in developing various intervention scenarios and cost benefit calculations – irrespective of the final political decision made. We don't claim that models make decisions, but that they are implicated (sometimes more sometimes less) in these decisions. What interests us here is that the original cost benefit model was challenged by an alternative model, even if this didn't suggest the interventions that were ultimately made (so it's more about the political power of the first model being challenged here).

**Reviewer:** *L459. "they conclude that models assuming that residents are well informed and have shared understandings of the water supply system might lead to an oversimplification of sociohydrological dynamics in a given location, and that more local involvement could mitigate this". OK (though not clear what type of quantitative model you might mean here) but is that just not another example of poor assumptions/poor practice/not fit-for-purpose/design for vested interest – ie. there is again surely a need to distinguish between the model and the way in which it is used.*

**Answer:** Comments above apply. It's about the encoding of assumptions in models and the obscuring of the encoder.

**Reviewer:** *L477 "which raises questions about the responsibility and accountability of those making and using models" Well yes, and that is the problem I have with most of your discussion since you are placing responsibility on the "model" and not on those who use and misuse them (not necessarily even the modeller, but even more so those who commission studies and/or use the results for decision making as you have demonstrated).*

*L606 "how quantitative models may help to foster transformative pathways towards more just and equitable water distributions." But why put the focus on the model here? What is needed is the political will for more just and equitable water distribution. Given that will, everything else would probably follow, but I do not see how you expect models (or modellers come to that) to influence the neoliberal capitalism or centralised communist systems that prevail in most countries where sustainable and equitable water issues are important. Look at the UK – we do not have the problems of degrading water quality because of any modelling or modeller issues. And announced this week is a relaxation of the rules on new housing developments in respect of water quality. A relaxation of rules in designating water quality categories is also expected now that we do not have to conform to the EU WFD standards.*

*Underlying many of the examples you provide is exactly that political framework, outside the control of modellers and their models.*

**Answer:** As explained above we see the model as being political as its characteristics and outcomes enable and constrain uses. It is the interaction between the commissioners, model, model makers and users and stakeholders that leads to policy decisions and implementation. In the case of the relaxation of water quality rules, there are many groups that will engage with trying to protect water quality in the UK. Modelling may play a role in this too.

Our call for power sensitive modelling is indeed a call for those who are part of modelling processes – including commissioners, modellers, and users, to become more critical or sensitive but not only to the context within which they use their models, but also through an understanding that there is an interaction between the way a model is made and how it functions in that context. It is therefore that we think that a power-sensitive approach could be useful.

*Reviewer: So while I appreciate the sentiments that lie behind the paper, I think you have got the framing wrong. Most of the power issues in moving towards better water management have very little to do with models or modellers – they are political (in the sense of the authors' use). So to say that models need to become power sensitive is not correct. It might be better put that modellers need to be more critical or sensitive to the context within which they use their models – both in terms of the framing of a project and whether a model is fit-for-purpose within such a project, but it is often the case that there is much that is outside their control (the legal basis of water rights in a country, the details of a project commission, etc). You give examples of where model outputs were rejected because they were in conflict with a desired outcome – that is surely a bigger problem of power.*

**Answer:** Yes, but important to us is again the encoding of power sensitivity in models (which might require ensembles of models). We shall try and make that clearer and that we indeed argue for a reconnection of the responsibility of model developers and users to models.

*Reviewer: So my apologies for these extended comments but it is something I have thought about over a long period of time. I think there are ways ahead (I have suggested some partial solutions in the past such as the condition tree / audit trail of assumptions, models of everywhere as ways of facilitating interaction and criticism, better evaluation of fitness-for-purpose and consideration of model uncertainties) but I think this paper needs to be reformulated much more about power in the political framing of the modelling process than assigning the responsibility to power sensitive models.*

**Answer:** We are grateful for this interaction which we believe will result in a refinement of our arguments as detailed above.

*Reviewer: Some other points*

*The choice of papers analysed seems incomplete. For the once case study that I know something about (Pickering) some papers are included (the papers by Lane et al. and Landström et al), but others are not (particularly Lane et al. 2011, Doing flood risk science differently in TIBG, and the papers on Pickering by Sarah Whatmore). Were these considered as having too much overlap or is it an indication that the methodology could have retrieved other relevant papers?*

**Answer:** In this case, it was the overlap between those papers that prevented us from citing all of them. As we aim to develop a growing database, we will make sure all papers on this project are included in the analysis.

*Reviewer: The Morgan and Morrison 1999 paper does not appear in the references.*

**Answer:** Thank you for pointing this out. We have added it.

*Reviewer: The Pielke Jr. reference appears twice.*

**Answer:** Thank you for pointing this out, we have deleted one reference.

**Answer to reviewer 2**

1       My apologies that this review is a little late. This paper needed quite some thought. I find its goals useful and there is some very useful material therein that the hydrological modelling community should be aware of. I should be honest about a few of my own positions with respect to this work. I don't particularly like this kind of methodology – finding literature this way often means that very interesting papers that perhaps frame themselves differently get lost – and because they are not found other kinds of framing that exist are left hidden. Second, I apologise that there is quite some reference to my own work in my report but I have been working on this topic for almost 20 years. Some of the ideas the authors have could and have been developed a lot more (and the work I refer to, notably Lane 2012, might help them access the work of others besides me). In especially the Discussion, I felt that the authors did not really show as much awareness as might be ideal of some of the dimensions that they address (mental images, framing) and which have already been discussed in relation to hydrology. There could be a lot more useful Discussion and crucially, perhaps my key recommendation, some much deeper thinking regarding what a power-sensitive modelling approach might look like and how it might come about (which itself would have to be a political act).

The detailed comments identify some of these issues that could be developed. As major comments though, and in addition, I think the following need some treatment in the paper.

Answer      Thank you for your supportive comments in relation to the aims of the review.

Regarding the methodology, we are aware we can never present a complete overview of all papers that engage with the influence of modelling in water, and reflect on this in the methodology as well. It is also challenging that different disciplines write about the topic with different jargons, and also that reflections on the influence of models and modelling are sometimes hidden in papers. We have therefore decided to make a start with collecting papers, combining two methods; a narrative review in which we identified papers that we were already aware of, and in addition did a systematic review to identify others. Also, we specifically have opted for HESS that has an open review process to invite people to share more papers.

We will enrich the discussion with references to papers that have addressed the question of what power-sensitive modelling might look like.

2       1. Models are themselves political objects – they contain intellectual property that can come to have commercial value – the history of how flood inundation modelling went from 1D to 2D in the UK was about primarily the protection of vested interests in intellectual property (for example, see one of the sections in Lane, S.N., November, V., Landström, C. and Whatmore, S.J., 2013. Explaining rapid transitions in the practice of flood risk management. Annals of the Association of American Geographers, 103, 330-42) – this is similarly reflected in James Porter's fascinating work about how flood maps get made (which is essentially based upon flood inundation modelling) and the negotiation between modellers/mappers and regulatory agencies (Porter, J. and Demeritt, D., 2012. Flood Risk Management, Mapping and Planning: The Institutional Politics of Decision-Support in England. Environment and Planning

Answer      Indeed, the economic dynamic of modelling has not been explicitly touched upon in the articles reviewed, and we acknowledge it is an important dynamic/driver in relation to the influence of models. We will update the review based on additional articles we have received and that will include a mention of this either as a separate section or in the comments, based on the outcomes of the review.

2A     A). There is a tendency to reduce politically sensitive modelling to the interface between models, modellers and society – and to overlook the politics of modelling itself and how this also needs to become

more politics sensitive. In turn this would sit this paper also in wider Science-Technology-Studies perspectives on power within scientific practice. Politics exists within modelling communities and this can have a profound impact on how modelling is done. Politics can constrain what is modelled and how. This needs more consideration and discussion.

Answer Thank you for bringing this up. The politics of science in relation to models has been included in the review, especially in sections 4.1 and 4.2 on 'Mental models and policy projects', and 'The influence of modellers' choices'. Where possible, we will elaborate based on the additional papers we will add in the review. Also, we agree it can be more specific in the discussion, as well as in the call for power-sensitive modelling and will update the article accordingly.

3    2. If we accept the hypothesis that power is imbricated with and within models of the water environment (which I do), the question becomes what to do. The paper has very little on what this means for practice. A starting point may be the obvious point that we should not make recourse to some kind of naïve goal that we should exclude power and politics from modelling such that models give us access to some kind of higher, truer knowledge and models can then reign supreme ; that would simply be an internal contradiction in that it would imply the transfer of power to models and modellers (that is itself be a political statement). So, if power and politics cannot be excluded what would a more democratic form of modelling mean in practice. What do we mean by "democratic"? How might different kinds of politics (e.g. majoritarian versus minoritarian) lead to models being used, power-sensitively, in different ways? What are the ways of using models that allow those excluded from decision making to become included in decision-making through the use of models. When I got to the end of the discussion I felt short-changed that after all the examples this hadn't been thought through.

Answer Thank you for bringing up the question: what now? This paper is part of ongoing efforts to understand the power related to modelling, and other articles are planned that engage with the question: how do we practically engage with the power of models? We base this work on an understanding that each model and modelling process is different and that this requires a flexible approach.

What we will do in the current article is to provide a further explanation in relation to the four elements (the mental models and policy projects, the influence of modellers choices, the impacts models have, and engaging with non-modellers) more explicitly, as well as reflect more on practical examples from the reviewed papers on how to engage in a power-sensitive way.

4    3. When I wrote the 2014 paper published in HESS I should confess that this was motivated by an extreme skepticism (that I still hold strongly) of socio-hydrology and the focus of the Panta Rhea decade; this included the appearance of diagrams coupling models of people to models of the hydrological cycle that reminded me of those the Dick Chorley and others were advocating in the 1960s and 1970s. Yet again, we as hydrologists seemed to be bolting people onto the natural environment with very little clue as to actually what the "socio-" really means. This was ironic because at the same time there was much interest from social scientists, notably Linton, Budd etc., in the Hydrosocial cycle which had, in my view, a much deeper understanding of how people and water are truly coupled. Now, I regret leaving in Lane (2014) this argument a little too subtle. But, I think the distinction between socio-hydrology and the Hydrosocial cycle does actually matter to this article because the need for political sensitivity reflects (for ethical reasons, at the very least) the need for a much more nuanced understanding of how hydrological models fit into the world. This is lost at the moment in this article. It probably does not need much discussion added but the notion of a hydrosocial cycle and the mutual influences between people and water strengthen the argument that we should also take a political perspective for how hydrological modelling is done.

Answer Thank you for bringing this up, and we fully agree with the importance of socio-political variables in understanding water systems. However, we did not explicitly focus on comparing socio-hydrology and hydrosocial studies because that could (and perhaps should) very well be a new paper. Nevertheless, we acknowledge that these different ontologies are important to mention, not in the least as the hydrosocial cycle acknowledges the political as a transformational force capable of changing landscapes. We will dedicate time to engage with these different ontologies in the introduction and discussion.

*Minor comments*

5     L36 – but also by the situated nature of the modeller themselves – and the work they do to come to trust their own models – see Lane, S.N., 2012. Making mathematical models perform in geographical space(s). Chapter 17 in Agnew, J. and Livingstone, D. Handbook of Geographical Knowledge. Sage, London
This is a very good point; we will integrate this in the article.

6     L36-7 – your references here, unless I am mistaken, largely support the counter-factual – do you have any better references of modellers themselves presenting their models as neutral tools? Without this there is a risk that you are setting up a "straw doll".

*Answer* — Indeed, often these positions are not made explicit, which is also shown through the literature review we did. Neutrality is often assumed and non-neutrality needs to be actively proven to challenge the dominant paradigm. Hence the chosen references.

7     L37-8 – I don't agree with this as it is written – we may think models can travel easily between places but much modelling in practice is about making them work in particular settings. There are studies of hydrological modelling in practice – see for instance Landstrom, C., Whatmore, S.J. and Lane, S.N., 2011. Virtual Engineering: computer simulation modelling for UK flood risk management. Science Studies, 24, 3-22

*Answer* — A challenge in our article is that we cover many different models, at different moments of their development. Some require a lot of work to make them fit, and others travel more easily based on how they are designed and what they do. We will tweak the sentence: "An additional challenge is that models become increasingly complex and some travel easily between places of application."

8     L63 – the notion of power sensitive modelling is interesting – but I think the paper is missing a wider link here into science technology studies that argues that power is an inevitable (and sometimes malign) component of any piece of scientific investigation. Isabelle Stengers has treated this generally (her book "Une autre science est possible" is probably the simplest entry point); there is then work around any one of a number of disciplines exploring this as well as some very specific to modelling of the environment (Demeritt, 2001 is a classic in this sense). I appreciate that the paper is about water but it is not well situated with respect to the wider STS work on the demonstrated influence of power upon scientific practice (and the very obvious point that the moment you advocate a model as neutral and therefore *the* valid basis for decision-making, you immediately transfer power to the model and the modeler, that is you make a political decision). Some of these examples also cloud be used to develop the argument around L90 a bit more – where it is under-developed in terms of what we already know about the relationship between power and modelling.

*Answer* — This comment strongly relates to the second point you make (or 1A), and we agree that, through the review, we have focused mostly on the influence of models in practice, and less on power in scientific practice. We will bring this out more explicit in section 2, defining models, as well in the discussion, and also bringing out these points through the STS work we have included, such as Knorr-Cetina, 1999, Bijker, 2017; Bijker et al., 1987; Latour, 2000; Latour and Woolgar, 1986; MacKenzie and Wajcman, 1999.

9     Shackley, S., Risbey, J., Stone, P., and Wynne, B.: Adjusting to policy expectations in climate change modeling – An interdisciplinary study of flux adjustments in coupled atmosphere-ocean general circulation models. Climatic Change, 43, 413–454, 1999
Demeritt, D.: The Construction of Global Warming and the Politics of Science, Annals of the Association of American Geographer, 91, 307–337, 2001
Lahsen, M.: Seductive simulations? Uncertainty distribution around climate models, Social Studies of Science, 35, 895–922, 200B.

Sundberg, M.: The Everyday World of Simulation Modeling: The Development of Parameterizations in Meteorology, Science Technology and Human Values, 34, 162–181, 2009
Brysse, K., Oreskes, N., O'Reilly, J., and Oppenheimer, M.: Climate change prediction: Erring on the side of least drama?, Global Environmental Change, 23, 327–337, 2013.

Answer

Thank you for these references. We had initially chosen to keep the introduction and definition of models short, and focus on showcasing and learning from the "water"-articles to acknowledge the specific approaches and culture that define how models are used and discussed in the water sector. We also see the importance of bringing in lessons from other fields, and will dedicate space for this in the discussion.

10    L84-88 See Lane (2012) op cit.

Answer

Thank you for the reference.

11    L108 – water is not just about hydrology but also hydraulics – is there a reason you did not also look at hydraulics. Note also that this way of searching likely misses more specific model applications or where models are not really referred to as "models" even if modelling is an integral part of what is being done.

Answer

This is true, and we will redo the query to include hydraulic models and other types (hydro-dynamic for instance) by adapting the query to:
TITLE-ABS-KEY ( "water model*" OR "hydr* model*" ) AND TITLE-ABS-KEY ( justice OR equit* OR politic* OR ethic* ).
(including 307 articles instead of 293, part of which can be explained by including now the 2nd half of 2022. We will also include the articles published in 2023, resulting in 323 articles included in the systematic review)

12    L115 – but isn't this also an important result in the context of this paper?

Answer

Yes, it is indeed an extremely important outcome! There is very little written, unfortunately. We will highlight this more explicitly in the article as outcome.

13    L124 – clarify what you mean by "narrative style"

Answer

We have changed this to 'narrative review' to clarify

**Answer to the reviewer 3**

**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.

*[handwritten margin note: this could mean that citations become crucial]*

**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.

*[handwritten margin note: still not clear]*

**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,

*[handwritten note at bottom: yes I know 14 these terms. still vague]*

*How?* [handwritten annotation]

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'.

*[handwritten margin note: The answer to tunnel vision is not interdis perse. (It can be a non tunnel vision)]*

**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.

The discussion and conclusion are very similar. I still think the paper lacks an overall linking of the themes. It still seems to have no other conclusion beyond "these issues matter." Fair enough, but that was already the starting point.